

SEQUENCE LISTING

(1) GENERAL INFORMATION:

- (i) APPLICANT: HUSE, WILLIAM D.
- (ii) TITLE OF INVENTION: SURFACE EXPRESSION LIBRARIES OF HETEROMERIC RECEPTORS
- (iii) NUMBER OF SEQUENCES: 75
- (iv) CORRESPONDENCE ADDRESS:
 (A) ADDRESSEE: PRETTY, SCHROEDER, BRUEGGEMANN & CLARK
 (B) STREET: 444 SO. FLOWER STREET, SUITE 200
 (C) CITY: LOS ANGELES
 (D) STATE: CALIFORNIA
 (E) COUNTRY: UNITED STATES
 (F) ZIP: 90071
- (v) COMPUTER READABLE FORM:
 (A) MEDIUM TYPE: Floppy disk
 (B) COMPUTER: IBM PC compatible
 (C) OPERATING SYSTEM: PC-DOS/MS-DOS
 (D) SOFTWARE: PatentIn Release #1.0, Version #1.25
- (vi) CURRENT APPLICATION DATA:
 (A) APPLICATION NUMBER:
 (B) FILING DATE:
 (C) CLASSIFICATION:
- (vii) ATTORNEY/AGENT INFORMATION:
 (A) NAME: CAMPBELL, CATHRYN A.
 (B) REGISTRATION NUMBER: 31,615
 (C) REFERENCE/DOCKET NUMBER: F31 8882
- (ix) TELECOMMUNICATION INFORMATION:
 (A) TELEPHONE: 619-535-9001
 (B) TELEFAX: 619-535-8949

(2) INFORMATION FOR SEQ ID NO:1:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 7445 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: both
 (D) TOPOLOGY: circular

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

AATGCTACTA CTATTAGTAG AATTGATGCC ACCTTTTCAG CTCGGCGCCG AAATGAAAAAT 60
 ATAGCTAAAC AGGTTATTGA GCATTTCGGA AATGTATCTA ATGGTCAAAAC TAAATCTACT 120
 CGTTCGCAGA ATTGGGAATC AACTGTIACA TGGAAAGAAA CTTCACAGACA CGGTACTITTA 180
 GTTGCATATT TAAAACATGT TCAGCTACAG CACCAGATTC AGCAATTAAG CTGTAAGCCA 240
 TCTGCAAAAA TGACCTCTTA TCAAAAGGAG CAATTAAGG TACTCTCTAA TCCTGACCTG 300
 TTGGAGTTTG CTTCGGGTCT GGTTCGCTTT GAAGCTCGAA TTAACACGGG ATATTGGAAG 360
 TCTTTCGGGC TTCTCTCTAA TCTTTTIGAT GCAATCCGCT TTGCTTCTGA CTATAAAGT 420

CAGGGTAAAG	ACCTGATTTT	TGATTTATGG	TCATTCTCGT	TTTCTGAACT	GTTTAAAGCA	480
TTTGAGGGGG	ATTCAATGAA	TAITTTATGAC	GATTCCGCAG	TATTGGACGC	TATCCAGTCT	540
AAACATTTTA	CTATTACCCC	CTCTGGCAAA	ACTTCTTTTG	CAAAAGCCTC	TCGCTATTTT	600
GGTTTTTATC	GTGCTCTGGT	AAACGAGGGT	TATGATAGTG	TGCTCTTAC	TATGCTCGT	660
AATTCCTTTT	GGCGTTATGT	ATCTGCATTA	GTTGAATGCT	GTAITTCCTAA	ATCTCAACTG	720
ATGAATCTTT	CTACCTGTAA	TAATGTTGTT	CGGTTAGTTC	GTTTTATTTA	CGTAGATTTT	780
TCITCCCAAC	GTCCGTACTG	GTATAATGAG	CCAGTCTCTA	AAATGCCATA	AGGTAATTCA	840
CAATGATTAA	AGTTGAAATT	AAAGCATCTC	AAGCCCAATT	TACTACTCGT	TCTGGTGTTT	900
CTCGTCAGGG	CAAGCCTTAT	TCACTGAATG	AGCAGCTTTG	TTACGTTGAT	TTGGGTAATG	960
AATATCCGGT	TCTTGTCAAG	ATTACTGTTG	ATGAAGGTCA	GCCAGCCTAT	GCGCCTGGTC	1020
TGTACACCGT	TCATCTGTCC	TCITTCAAAG	TTGGTCAGTT	CGGTTCCGCT	ATGATTGACC	1080
GTCTGCGCCT	CGTTCGGGCT	AAGTAACATG	GAGCAGGTGG	CGGATTTCTA	CACAATTTAT	1140
CAGGCGATGA	TACAAATCTC	CGTTGTACTT	GTITTCGGCG	TTGGTATAAT	CGCTGGGGGT	1200
CAAAGATGAG	TGTTTTAGTG	TATTCCTTGG	CGCTTTCGTT	TTTAGGTTGG	TGCCTTCGTA	1260
GTGGCATTAC	GTATTTTACC	CGTTTAATGG	AAACTTCCTC	ATGAAAAAGT	CTTTAGTCTT	1320
CAAAGCCTGT	GTAGCCGTTG	CTACCTCTGT	TCCGATGCTG	TCTTTCGCTG	CTGACGGTGA	1380
CGATCCCGCA	AAAGCGGCTT	TAACTCCCT	GCAAGCTCA	GCGACCGAAT	ATATCGGTTA	1440
TCCGTGGGCG	ATGGTTGTTG	TCATTGTGCG	CGCAACTATC	GGTATCAAGC	TGTTTAAGAA	1500
ATTCACCTCG	AAGCAAGCT	GATAAACCGA	TACAATTAAA	GGCTCCTTTT	GGAGCCTTTT	1560
TTTTTGAGA	TTTTCAACGT	GAATAAAATTA	TTATTCGCAA	TTCTTTTAGT	TGTTCTTTTC	1620
TAITCTCACT	CCGCTGAAAC	TGTTGAAAGT	TGTTTAGCAA	AACCCCATAC	AGAAAAATTCA	1680
TTTACTAACG	TCGTGAAAGA	CGACAAAAC	TTAGATCGTT	ACGCTAACTA	TGAGGGTTGT	1740
CTGTGGAATG	CTACAGGCGT	TGTAGTTTGT	ACTGGTGACG	AAACTCAGTG	TTACGGTACA	1800
TGGGTTCTTA	TTGGGCTTGC	TATCCCTGAA	AATGAGGGTG	GTGGCTCTGA	GGGTGGGGGT	1860
TCTGAGGGTG	GGGGTTCTGA	GGGTGGGGGT	ACTAAACCTC	CTGAGTACGG	TGATACAGCT	1920
ATTCGGGGGT	ATACTTATAT	CAACCTCTCT	GACGGCACTT	ATCCGCTCTG	TACTGAGCAA	1980
AACCCCGCTA	ATCCTAATCC	TTCTCTTGAG	GAGTCTCAGC	CTCTTAATAC	TTTCAATGTT	2040
CAGAATAATA	CGTTCGAAAA	TAGGCAGGGG	GCATTAAGTG	TTTATACGGG	CAGCTGTIAGT	2100
CAAGGCACATG	ACCCCGTTAA	AACCTTATAC	CAGTACATCT	CTGTATCATG	AAAGGCCATG	2160
TATGAGCGCT	ACTGGAACGG	TAAATTGAGA	GACTGGGCTT	TCCATTCTGG	GTTTAATGAA	2220
GATCCATTCT	TTTGTAATA	TCAAAGGCCAA	TGCTCTGACC	TGCTCAACCC	TGCTGTCAAT	2280
GCTGGCGGGG	GCTCTGGTGG	TGGTCTGGT	GGCGGCTCTG	AGGGTGGTGG	CTCTGAGGGT	2340
GGCGGTTCTG	AGGGTGGCGG	CTCTGAGGGA	GGCGGTTCCG	GTGGTGGCTC	TGGTCCGGGT	2400
GATTTTGATT	ATGAAAAGAT	GGCAAAACGT	AATAAGGGGG	CTATGACCGA	AAATGCGGAT	2460

GAAGACGGCG	TACAGTCTGA	CGCTAAAGGG	AAACTTGATT	CTGTCGGTAC	TGATTACGGT	2520
GCTGCTATCG	ATGGTTTCAAT	TGGTGACGTT	TCCGGCCTTG	CTAATGGTAA	TGGTGCTACT	2580
GGTGATTTTG	CTGGCTCTAA	TTCCCAAATG	GCTCAAGTCG	GTGACGGTGA	TAATTACGCT	2640
TTAATGAATA	ATTTCCGTCA	ATATTACCT	TCCCTCCGTC	AATCGGTTGA	ATGTCGGCCT	2700
TTTGCTTTA	GCGCTGGTAA	AGCATATGAA	TTTTCTATTG	ATTGTGACAA	AATAAACTTA	2760
TTCCGTGGTG	TCTTTGGCTT	TCTTTTATAT	GTGCGCACCT	TTATGTATGT	ATTTTCTACG	2820
TTTGCTAACA	TACTGGCTAA	TAAGGAGTCT	TAATCATGCC	AGTTCCTTTG	GGTATCCGCT	2880
TATTATTGGG	TTTCTCGGT	TTCTTCTGG	TAACTTTGTT	CGGCTATCTG	CTTACTTTTC	2940
TTAAAAAGGG	CTTCGGTAAG	ATAGCTATTG	CTATTTCAAT	GTTTGTGCT	CTTATTATTG	3000
GGCTTAACTC	AATTCCTGTG	GGTTATCTCT	CTGATAATTAG	CGCTCAATTA	CCCTCTGACT	3060
TTGTCAGGG	TGTTCACTTA	ATCTCTCCGT	CTAATGGGCT	TCCCTGTTTT	TATGTTATTG	3120
TCTCTGAAA	GGCTGCTATT	TTCATTTTTG	ACGTTAAACA	AAAAATCGTT	TCTTATTGCG	3180
ATTGGGATAA	ATAATATGGC	TGTTTTATTT	CTAAGTGGCA	AATTAGGCTC	TGGAAAGACG	3240
CTGCTTAGCG	TTGGTAAGAT	TCAGGATAAA	ATTCTAGCTG	GGTGCAAAAT	AGCAACTAAT	3300
CTTGATTTAA	GGCTTCAAAA	CCTCCCGCAA	CTCGGAGGT	TGGCTAAAC	GGCTCGCGTT	3360
CTTAGAATAC	CGGATAAGCC	TTCTATATCT	GAITTCCTTG	CTATTGGGCG	CGGTAATGAT	3420
TCCTACGATC	AAAATAAAAA	CGGCTTGCCT	GTCTCCATC	ACTCGCGTAC	TTGCTTTAAT	3480
ACCGGTTCTT	GGAATGATAA	GGAAGACAG	CGGATTATTC	ATTGGTTTCT	ACATGCTCGT	3540
AAATTAGGAT	GGGATATTAT	TTTTCTGTT	CAGGACTTAT	CTATTGTTGA	TAAACAGGCG	3600
CGTTCTGCAT	TAGCTGAACA	TGTTGTTTAT	TGTCGTCGTC	TGGACAGAAT	TACTTTACCT	3660
TTTGCGGTA	CTTTATATTG	TCTTATTAAT	GGCTCGAAAA	TGCTCTGCGC	TAAATTACAT	3720
GTGCGGTTG	TAAATATGCG	CGATTCTCAA	TAAAGCCCTA	CTGTTGAGCG	TTGCGTTTAT	3780
ACTGTAAGA	ATTGTATATA	CGCATATGAT	ACTAAACAGG	CTTTTCTCTAG	TAATTATGAT	3840
TCCGGTGTTT	ATTCTTATTT	AACGCCCTTAT	TTATCACACG	GTCCGTTATTT	CAAAACCATTA	3900
AATTTAGGTC	AGAAGATGAA	GCTTACTAAA	ATATATTGGA	AAAAGTTTTC	ACCGGTTCTT	3960
TGCTTCCGA	TTGGATTTCG	ATCAGCAATT	ACATATAGTT	ATATAACCCA	ACCTAAGCCG	4020
GAGGTAAAA	AGGTAGTCTC	TCACACCTAT	GAITTTGATA	AATTCACTAT	TGACTTTTCT	4080
CACGGTGTTA	ATCTAAGCTA	TGCTATGTTT	TTCAAGGATT	CTAAGGGAAA	ATTAATTAAT	4140
AGCGAGGATT	TACAGAAGCA	AGGTTATTGA	CTCACATATA	TTGATTATAT	TAGTGTTTCC	4200
ATTAAAAAAG	GTAATTCAAA	TGAAATTGTT	AAATGTAATT	AATTTTGTTT	TCTTATGTTT	4260
TGTTTCATCA	TCTTCTTTTG	CTCAGGTAAT	TGAAATCAAT	AATTCGCGTC	TGCGCGAATT	4320
TGTAAGTTGG	TATTCAAAGC	AATCAGGCGA	ATCCGTTAAT	GTTTCTCCCG	ATGTAAAAAG	4380
TACTGTTACT	GTATATTCAAT	CTGACGTTAA	ACCTGAAAAAT	CTACGCAATT	TCTTTATTTC	4440
TGTTTACGCT	GCTAATAAAT	TTGATATGGT	TGTTCAAATT	CCTTCCTATA	TTCAAGAGTA	4500

TAATCGAAAC AATCAGGATT ATATTGATGA ATGGCATCA TCTGATAATC AGGAATATGA	4560
TGATAATTCC GCTCTCTCTG GTGGTTTCTT TGTTCGGCAA AATGATAATG TTAATAAAC	4620
TTTTAAAAAT AATAACGCTC GGGCAAAGGA TTAAATACGA CTGTCCGAAT TGTITGTA	4680
GTCTAATACT TCTAAATCCT CAAATGTATT ATCTATTGAC GGCTCTAATC TATTAGTTGT	4740
TAGTGCACCT AAAGATATTG TAGATAACCT TCTCAATTG CTCTCTACTG TTGATTGGCC	4800
AAGTGCACAG ATATTGATTG AGGGTTTGAT ATTTGAGGTT CAGCAAGGTG ATGCTTTAGA	4860
TTTTTCATTG CTGCTGGCTG CTCAGCGTGG CACTGTGCGA GCGCGTGTTA ATACTGACCG	4920
CCTCACCTCT GTTTTATCTT CTGCTGGTGG TTCGTTCGGT ATTTTAAATG CCGATGTTTT	4980
AGGGCTATCA GTTCGGCGAT TAAAGACTAA TAGCCATTCA AAAATATTGT CTGTGCCACG	5040
TATCTTACG CTTCAGGTC AGAAGCGTTC TATCTCTGT GGCAGAAATG TCGCTTTTAT	5100
TACTGCTGTG GTGACTGGTG AATCTGCGAA TGTAAATAAT CCATTCAGA CGATTGACCG	5160
TCAAAATGTA GGTATTTCCA TGAGCGTTCT TCTCTTGGA ATGGCTGGCG GTAATATTCT	5220
TCTGATATT ACCAGCAAGG CCGATAGTTT GAGTCTCTCT ACTCAGGCAA GTGATGTTAT	5280
TACTAATCAA AGAAGTATTG CTACAACCGT TAAATTGGCT GATGGACAGA CTCTTTACT	5340
CGGTGGCTTC ACTGATTATA AAAACACTTC TCAAGATTCT GCGGTACCGT TCTGTCTAA	5400
AATCCCTTAA ATCGGCTCC TGTTTAGCTC CCGCTCTGAT TCCAACGAGG AAAGCACGTT	5460
ATACGTGCTG GTCAAAGCAA CCAATAGTACG GCGCGTGTAG CCGCGCATTG AGCGCGCGGG	5520
GTGTGGTGGT TACGCGCAGC GTGACCGCTA CACTTGCGAG CCGCCTAGCG CCGCTCCTT	5580
TGCTTTCTT CCCTTCCTTT CTGCGCACGT TCGCGCGGTT TCGCGGTCAA GCTCTAAATC	5640
GGGGGCTCCC TTTAGGGTTC CGATTTAGTG CTTTACGGCA CCTCGACCCC AAAAACTTG	5700
ATTTGGGTGA TGCTTCACGT AGTGGGCGAT CCGCGTGATA GACGGTTTTT CGCCCTTGA	5760
CGTTGGAGTC CAGCTTCTTT AATAGTGGAC TCTTGTCCA AACTGGAACA ACACTCAACC	5820
CTATCTCGGG CTATCTTTTT GATTTATAAG GGATTTTCCC GATTTCGGAA CCACCATCAA	5880
ACAGGATTTT CGCTGCTGG GGCAAAACAG CGTGGACCGC TTGCTCCAACT TCTCTCAGGG	5940
CGAGGCGGTG AAGGCGAATG AGCTGTTGCC CGTCTCGCTG GTGAAAAGAA AAACCAACCT	6000
GGCGCCCAAT ACGCAAAACCG CCTCTCCCGG GCGGTGCGCG GATTCTATTA TGCAGTGGC	6060
ACGACAGGTT TCCGACTGAG AAAGCGGGCA GTGACGCGAA CGCAATTAAAT GTGAGTTAGC	6120
TCACTCATTA GGCACCCGAG GCTTTACACT TTATGCTTCC GGCTCGTATG TTGTGTGGAA	6180
TTGTGAGCGG ATAACAATTT CACACCGCTC ACTTGGCACT GCGCGTCTTT TTACAACGTC	6240
GTGACTGGGA AAACCTGGC GTTACCCAAG CTTTGTACAT GGAGAAAATA AAGTGAACA	6300
AAGCACTAAT GCATGCGAC TCTTACCGTT ACCGTTACTG TTTACCCGTG TGACAAAAGC	6360
CGCCACGTC CAGCTGCTCG AGTCAGGCGT ATTGTGCCCC GGGGATTGTA CTAGTGGATC	6420
CTAGGCTGAA GGGCATGACC CTGCTAAGGC TGCATTCAAT AGTTACAGG CAAGTCTTAC	6480
TGAGTACATT GGCTACGCTT GGGCTATGCT AGTATGTATA GTTGTGCTA CCATAGGGAT	6540

TAAATTATTC AAAAAGTTTA CGAGCAAGGC TTCTTAAGCA ATAGCGAAGA GGCCCGCACC 6600
 GATCGCCCTT CCCACAATT GCGCAGCGCT AATGGCGAAT GCGCCTTTGC CTGGTTTCCG 6660
 GCACCAGAAG CCGTGCCGGA AAGCTGGCTG GAGTGGCATC TTCCTGAGGC CGATACGGTC 6720
 GTGCTCCCTT CAACTGGCA GATGCACGGT TACGATCGGC CCATCTACAC CAACGTAACC 6780
 TATCCCATTA CCGTCAATCC GCGGTTTCTT CCCACGGAGA ATCCGACGGG TTGTTACTCG 6840
 CTCACATTA ATGTTGATGA AAGCTGGCTA CAGGAAGGCC AGACGCGAAT TATTTTGTAT 6900
 GCGGCTCCTA TTGGTAAAAA AATGAGCTGA TTTAACAAAA ATTAAACGGC AATTTTAACA 6960
 AAATATTAAC GTTACAATT TAAATATTG CTTATACAAT CTTCCTGTTT TTGGGGCTTT 7020
 TCTGATTATC AACCGGGGTA CATATGATTG ACATGCTAGT TTAAQGATTA CCGTTCATCG 7080
 ATTCTCTGTG TTGCTCCAGA CTCTCAGCCA ATGACCTGAT AGCCTTTGTA GATCTCTCAA 7140
 AAATAGCTAC CCTCTCCGGC ATTAATTAT CAGCTAGAAC GGTGAATAT CATATTGATG 7200
 GTGATTGAC TGCTCCGGC CTCTCTCAGC CTTTTGAATC TTTACCTACA CATTACTCAG 7260
 GCATTGCATT TAAATATAT GAGGGTTCTA AAAATTTTIA TCCTTCCGTT GAAATAAAGG 7320
 CTCTCCGGC AAAAGTATTA CAGGGTCATA ATGTTTTTGG TACAACCGAT TTAGCTTTAT 7380
 GCTCTGAGGC TTTATTGCTT AATTTTGCTA ATGTTTGCC TTGCTGTAT GATTTATTGG 7440
 ACGTT 7445

(2) INFORMATION FOR SEQ ID NO:2:

- (1) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 7317 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: both
 - (D) TOPOLOGY: circular

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

AATGCTACTA CTATTAGTAG AATTGATGCC ACCTTTTCAG CTGCGCGCCC AAATGAAAT 60
 ATAGCTAAAC AGGTTATTGA CCATTTCGGA AATGTATCTA ATGCTCAAC TAAATCTACT 120
 CGTTCCGAGA ATTGGGAATC AACTGTTACA TGGAAAGAAA CTCCAGACA CCGTACTTTA 180
 GTTGCAATTT TAAACATGTT TGAGCTACAG CACCAGATTC AGCAATTAAG CTCTAAGCCA 240
 TCCGCAAAAA TGACCTCTTA TCAAAAGGAG CAATTAAAGG TACTCTCTAA TCCTGACCTG 300
 TTGAGTTTG CTTCGGGTCT GGTTCGGTTT GAAGCTCGAA TTAACCGCG ATATTGAAG 360
 TCTTTGGGC TTGCTCTTAA TCTTTTGTAT GCAATCCGCT TTGCTCTGTA CTATAATAGT 420
 CAGGCTAAGC ACCTGATTTT TGATTATGG TCATTCTGCT TTTCTGAAGT GTTTAAAGCA 480
 TTTGAGGGGG ATTCAATGAA TATTTATGAC GATTCCGCAG TATCGAGCG TATCCAGTCT 540
 AAACATTTTA CTATACCCC CTCTGGCAAA ACTTCCTTTG CAAAAGCTC TCGTATTTT 600
 GGTTTTATC GTGCTGTGTT AAACGAGGCT TATGATAGTG TTGCTCTTAC TATGCTCTGT 660
 AATTCTTTT GCGGTATGTT ATCTGCATTA GTTGAATGTG GTATTCTTAA ATGCTCAACTG 720

ATGAATCTTT	CTACCTGTAA	TAATGTTGTT	CCGTTAGTTC	GTTTTATTAA	CGTAGATTTT	780
TCTTCCCAAC	GCCTCGACTG	GTATAATGAG	CCAGTTCCTA	AAATCCGATA	AGGTAATTC	840
CAATGATTAA	AGTTGAAAT	AAACCATCTC	AAGCCCAATT	TACTACTCGT	TCIGGTGTTT	900
CTCGTCAGGG	CAAGCCTTAT	TCAGTGAATG	AGCAGCTTTG	TTACGTTGAT	TTGGGTAATG	960
AATATCCGGT	TCTTGTCAAG	ATTACTCTTG	ATGAAGGTCA	GCCAGCCTAT	CGGCTGGTC	1020
TGTACACCGT	TCATCTGTCC	TGTTTCAAAG	TTGGTCAGTT	CGGTTCCCTT	ATGATTGAAC	1080
GTCTGCGGCT	CGTTCCGGCT	AAGTAACATG	GAGCAGGTCC	CGGATTTCGA	CACAATTTAT	1140
CAGGCGATGA	TACAAATCTC	CGTTCTACTT	TGTTTCGGCC	TTGGTATAAT	CGCTGGGGGT	1200
CAAAGATGAG	TGTTTTAGTG	TATTCTTTCC	CCTCTTTCGT	TTAGGTTGG	TGCTTCGTA	1260
GTGGCATTAC	GTATTTTACC	CGTTTAAATG	AAACTTCCTC	ATGAAAAAGT	CTTTAGTCTT	1320
CAAAGCCTCT	GTAGCCGTTG	CTACCTCTGT	TCCGATGCTG	TCTTTCGGTG	CTGAGGGTGA	1380
CGATCCGCA	AAAGCGGCTT	TTAACTCCCT	GCAAGCCTCA	GCGACCGAAT	ATATCGGTTA	1440
TGCGTGGGG	ATGGTTCGTT	TCATTGTCCG	CGCACTATC	GCTATCAAGC	TGTTTAAGAA	1500
ATTCACTCG	AAAGCAAGCT	GATAAACCGA	TACAATTAAA	GGCTCCTTTT	GGAGCCTTTT	1560
TTTTTGAGA	TTTTCACGCT	GAAAAAATA	TTATTGCGAA	TTCTTTAGT	TGTTCTTTTC	1620
TATTCTCACT	CCGCTGAAAC	TGTTGAAAGT	TGTTTAGCAA	AACCCCATAC	AGAAAATTCA	1680
TTTACTAACC	TCTGGAAAGA	CGACAAAAC	TTAGATCGTT	ACGCTAACTA	TGAGGGTTGT	1740
CTGTGGAATG	CTACAGGCGT	TGTAGTTTGT	ACTGGTGACC	AAACTCAGTG	TTACGGTACA	1800
TGGGTTCTTA	TTGGGCTTGC	TATCCCTGAA	AATGAGGGTG	GTGGCTCTGA	GGGTGGCGGT	1860
TCTGAGGGTG	GGGTTCTGA	GGGTGGCGGT	ACTAAACCTC	CTGAGTACGG	TGATACACCT	1920
ATTCCGGGCT	ATACTTATAT	CAACCTCTC	GACGGCACTT	ATCCGCTCGG	TACTGAGCAA	1980
AACCCCGCTA	ATCTTAATCC	TTCTCTTGAG	GAGTCTCAGC	CTCTTAATAC	TTTATGTTT	2040
CAGAATAATA	GCTTCGAAA	TAGGCAGGGG	GCATTAACCT	TTTATACGGG	CAGTGTACT	2100
CAAGGCATG	ACCCCGTTAA	AACCTTATAC	CAGTACATC	CTGTATCATG	AAAAGCCATG	2160
TATGACGCTT	ACTGCAACGG	TAAATTGAGA	GACTGGGCTT	TGCATTCTGG	CTTTAATGAA	2220
GATCCATTCG	TTTGTGAATA	TCAAGGCCAA	TGCTCTGACC	TGCTCAACCC	TCCTGTCAAT	2280
GCTGGCGGG	GCTCTGGTGG	TGGTCTGGT	GGCGGCTCTG	AGGCTGGTGG	CTCTGAGGGT	2340
GGCGGTTCTG	AGGGTGGCGG	CTCTGAGGGA	GGCGGTTCCG	GTGGTGCTC	TGGTTCGGGT	2400
GATTTTGATT	ATGAAAAGAT	GGCAACGGT	AATAAGGGGG	CTATGACCGA	AAATGCGGAT	2460
GAAAACGGCG	TACAGTCTGA	CGCTAAAGGC	AAACTTGATT	CTGTGGCTAC	TGATTACGGT	2520
GCTGTATCC	ATGGTTTCAT	TGGTGACGTT	TCCGGCCTTG	CTAATGCTAA	TGGTGCTACT	2580
GGTGATTTTG	CTGGCTCTAA	TTCCCAAATG	GCTCAAGTCG	GTGACGGTGA	TAATTCACCT	2640
TTAATGAATA	ATTTCCGTCA	ATATTACCT	TCCCTCCCTC	AATCGGTTGA	ATGTGCGCCT	2700
TTTGTCTTTA	GGGCTGGTAA	ACCATATGAA	TTTTCTATTG	ATTGTGACAA	AATAAACTTA	2760

TTCCGTGGTG TCTTTGGCTT TCTTTTATAT GTTGCCACCT TTATGTATGT ATTTTCTACG	2820
TTTGTAACA TACTGGCTAA TAAGGAGTCT TAATCAIGCC AGTTCITTTG GGTATCCCGT	2880
TAITATTGGG TTCTCTGGT TTCTTCTGG TAACITTGTT CGGCTATCG CTACTTTTC	2940
TTAAAAAGG CTTCGGTAAAG ATAGCTAATG CTATTTCAAT GTTCTTGCT CTTATTAATG	3000
GGCTTAACTC AATTCTTGTG GGTATCTCT CTGATATTAG CGCTCAATTA CCTCTGACT	3060
TTGTTACGGG TGTTCAGTTA ATTCCTCCGT CTAATGGCGT TCCCTGTTTT TATGTTATFC	3120
TCTCTGAAA GGCTGCTATT TGCATTTTGG ACGTAAACA AAAAATCGTT TCTATTTTGG	3180
ATTGGGATAA ATAATATGGC TGTTTATTTT GTAACTGGCA AATTAGGCTC TCGAAAGACG	3240
CTCGTTAGCG TTGGTAAGAT TCAGGATAAA ATTGTAGCTG GGTGCAAAAT AGCAACTAAT	3300
CTTGATTAA GGCTTCAAAA CCTCCGCCAA GTCGGGAGGT TCGTAAAAAC GCCTCGCGTT	3360
CTTAGAATAC CGGATAAGCC TTCTATATCT GATTTGCTTG CTATTGGCGG CGGTAATGAT	3420
TCTACGATG AAAATAAAAA CGGCTTGGT GTTCTCGATG AGTGGCGTAC TTGGTTTAAT	3480
ACCGGTCTCT GGAATGATAA GGAAAGACAG GCGATTATTG ATTGGTTTCT ACATGCTCGT	3540
AAATTAGGAT GGGATATTAT TTTTCTGTT CAGGACTTAT CTATTGTTGA TAAACAGGGC	3600
CGTTCGCTAG TACGTGAACA TCTTCTTAT TCTCTGCTC TGCACAGAAT TACTTTACCT	3660
TTGTGCGTA CTTTATATTC TCTTATTACT GGCTCGAAAA TGGCTCTGCC TAAATTACAT	3720
GTGGCGTTG TTAATATGG CGATTCTCAA TTAAGCCCTA CTGTTGACGG TTGGCTTTAT	3780
ACTGGTAAGA ATTGTATATA GGCATATGAT ACTAAACAGG CTTTTCTAG TAATTATGAT	3840
TCCGGTGTTT ATCTTATTTT AACGCCTTAT TTATCACAGG GTCGGTATTT CAAACCAATTA	3900
AATTAGGTC AGAAGATGAA GCTTACTAAA ATATATTGTA AAAAGTTTT CACGGTTCTT	3960
TGCTTGGCA TTGGATTGTC ATCAGCATT ACATATAGTT ATATAACCCA ACCTAAGCCG	4020
GAGGTTAAAA AGGTAGTCTC TCAGACCTAT GATTTGATA AATTCACAT TGACTCTTCT	4080
CAGCGTCTTA ATCTAAGCTA TCGCTATGTT TTCAAGGATT CTAAGGGAAA ATTAATTAAT	4140
AGCGACGATT TACAGAAGCA AGGTTATTCA CTCACATATA TTGATTATG TACTGTTTCC	4200
ATTAAAAAAG GTAATTCAAA TGAATTTGTT AAATGTAATT AATTTTGTTT TCTTGATGTT	4260
TGTTTCATCA TCTTCTTTG CTCAGGTAAT TGAATGAAT AATTCGGCTC TCGCGGATT	4320
TGTAACTTGC TATTCAAAGC AATCAGGCGA ATCCGTTATT GTTCTCCCG ATGTAAAAAG	4380
TACTGTTACT GTATATTCAI CTCAGCTTAA ACCTGAAAA CTACGCAATT TCTTTATTTT	4440
TGTTTTACGT GCTAATAATT TTGATATGGT TGGTTCAATT CCTCCATAA TTCACAAGTA	4500
TAATCCAAAC AATCAGGATT ATATTGATGA ATTGCCATCA TCTATAATC AGGAATATGA	4560
TGATAAATTC GCTCCTTCTG GTGGTTTCTT TGTTCGGCAA AATGATAATG TTAAGTAAAC	4620
TTTTAAATTT AATAACGTTT GGGCAAAGGA TTTAATACGA GTTGTGAAT TGTTTGTAAA	4680
GTCTAATACT TCTAAATGCT CAAATGTATT ATCTATTGAC GGCTCTAATC TATTAGTTGT	4740
TAGTGCACCT AAAGATATTT TAGATAACCT TCCTCAATTC CTTTCTACTG TTGATTGGC	4800

AAC TGACCAG ATATTGATTC	AGGGTTTGAT	ATTGAGGTT	CAGCAAGGTT	ATGCTTTAGA	4860	
TTTTTCATTT	GCTGCTGGGT	CTCAGCGTGG	CAC TGTGCA	GCGGTTGTTA	ATACTGACCG	4920
CCTCACCTCT	GTTTTATCTT	CTGCTGGTGG	TTGTTGGGT	ATTTTTAATG	GCGATGTTTT	4980
AGGGCTATCA	GTTCGGCAT	TAAAGACTAA	TAGCCATTCA	AAAAATTGTT	CTGTGCCACG	5040
TATTCITACG	CTTTCAGGTC	AGAAGGGTTC	TATCTCTGTT	GCCGAGAATG	TCCCTTTTAT	5100
TACTGGTCTG	GTGACTGGTG	AATCTGCCAA	TGTAAATAAT	CGATTTCAGA	CGATTGAGCG	5160
TCAAAATGTA	GGTATTTCCA	TGAGCGTTTT	TGCTGTTGCA	ATGGCTGGCG	GTAATATTGT	5220
TCTGGATATT	ACCAGCAAGG	CGGATAGTTT	GAGTTCTTCT	ACTGAGGCAA	GTGATGTTAT	5280
TACTAATCAA	AGAAGTATTG	CTACAACGGT	TAATTGGGT	GATGGACAGA	CTCTTTTACT	5340
CGGTGGCCTC	ACTGATTATA	AAAAGACTTC	TCAAGATTCT	GCGGTACGGT	TCCTGTCTAA	5400
AATCCCTTAA	ATCGGCCTCC	TGTTTAGGTC	CCGCTCTGAT	TGCAACGAGG	AAAGCAGGTT	5460
ATACGTGCTC	GTCAAAGCAA	CCATAGTACG	CGCCCTGTAG	CGCGCATTA	ACCGCGGCGG	5520
GTGTGGTGGT	TACGCGCAGC	GTGACCGCTA	CACTTGCCAG	CGCCCTAGCG	CCCCTCCTT	5580
TGGCTTTCTT	CCCTTCCTTT	CTCGCCACGT	TGCGGGGCTT	TCCCGCTCAA	GCTCTAAATC	5640
GGGGGCTCCC	TTTAGGGTTC	CGATTTAGTG	CTTACGGCA	CCTCGACCCC	AAAAAACTTG	5700
ATTGGGTGTA	TGTTTCACGT	AGTGGGCCAT	CGCCCTGATA	GACGGTTTTT	CGCCCTTTGA	5760
CGTTGGAGTC	CAGGTTCITT	AATAGTGGAC	TCTTTTCCA	AACTGGAACA	ACACTCAACC	5820
CTATCTCGGG	CTATTCTTTT	GATTTATAAG	GGATTGTGCC	GATTTCGGAA	CCACCATCAA	5880
ACAGGATTTT	CGCCTGCTGG	GGCAAACGAG	CGTGGACCGC	TTGCTGCAAC	TCTCTCAGGG	5940
CCAGGCGGGT	AAGGGCAATC	AGCTGTGTGC	CGTCTCGCTG	GTGAAAAGAA	AAACCAACCT	6000
GGCGCCCAAT	ACGCAAAACG	CCTCTCCCCG	CGCGTTGGCG	GATTCTTAA	TGCAGCTGGC	6060
ACGACAGGTT	TCCCAGCTGG	AAAGCGGGCA	GTGAGCGCAA	CGCAATTAAT	GTGAGTTAGC	6120
TCACTGATTA	GGCACCCCGAG	GCTTTACACT	TTATGCTTCC	GGCTCGTATG	TTGTGTGGAA	6180
TTGTGAGCGG	ATAACAATTT	CACACGGCAA	GGAGACAGTC	ATAATGAAT	ACCTATTGCC	6240
TACGGCAGCG	GCTGGATTGT	TATTACTCGC	TGCCCAACCA	GCGATGGCGG	AGCTCGTGAT	6300
GACCGAGACT	CCAGATATCC	AACAGGAATG	AGTGTAAATT	CTAGACGGCG	TCAC TTGGGA	6360
CTGGCCGTGG	TTTTACAACG	TGTTGACTGG	GAAAACCCGT	GCGTTACCCA	AGCTTAATCG	6420
CCTTGCGAG	TTCCGTTTCG	CCAGCTGGCG	TAATAGCGAA	GAGGCGCGCA	CCGATCGCGC	6480
TTCCCAACAG	TTCCGCGAGC	TGAATGGCGA	ATGGCGCTTT	GCTCGCTTTC	CGGCACCAGA	6540
AGCGGTCCCG	GAAAGCTGGC	TGGAGTCCGA	TCTTCCTGAG	GCCGATACGG	TGTCGTCTCC	6600
CTCAAACTGG	CAGATGCACG	GTTACGATGC	GCCGATCTAC	ACCAAGCTAA	GCTATCCCAT	6660
TACGGTCAAT	CGCGGTTTG	TTCCCAACGA	GAATCCGACG	GCTTGTACT	CGCTCACATT	6720
TAATGTTGAT	GAAAGCTGGC	TACAGGAAGG	CCAGACCGGA	ATTATTTTTC	ATGGCGTTCC	6780
TATTGCTTAA	AAAATGAGCT	GATTTAACAA	AAATTTAACG	CGAATTTTAA	CAAAATATTA	6840

ACGTTTACAA TTTAAATATT TGCTTATACA ATCTTCCTGT TTTTGGGGCT TTCTGATTA	6900
TCAACCGGGG TACATATGAT TGACATGCTA GTTTTACGAT TACCGTTCAT CGATTCTCTT	6960
GTTTGCTCCA GACTCTCAGG CAATGACCTG ATAGCCTTTG TAGATCTCTC AAAAATAGCT	7020
ACCCTCTCCG GCATTAAATT ATCAGCTAGA ACGGTTGAAT ATCATATTGA TGGTGATTTC	7080
ACTGTCTCCG GCGTTTCTGA CCGTTTGAA TCTTTACCTA CAGATTACTC AGGCATTGCA	7140
TTTAAATAT ATGAGGGTTC TAAAAATTTT TATCCTTGGG TTGAAATAAA GCGTCTCCG	7200
GCAAAAGTAT TACAGGGTCA TAATGTTTTT GGTACAACCG ATTAGCTTT ATGCTCTGAG	7260
GCTTATTGC TTAATTTTGC TAATTCCTTG CCTTGCCTGT ATGATTIATT GGATGTT	7317

(2) INFORMATION FOR SEQ ID NO:3:

- (1) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 7729 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: both
 (D) TOPOLOGY: circular

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:3:

AATGCTACTA CTATTAGTAG AATTGATGCG ACCTTTTCAG CTCGCGCCCC AAATGAAAAT	60
ATAGCTAAAC AGGTTATTGA CCATTTCGGA AATGTATCTA ATGGTCAAAC TAAATCTACT	120
CGTTCGCAGA ATTGGGAATC AACTGTTACA TGGGAATGAAA CTTCAGACA CCGTACTTTA	180
GTTGCATATT TAAAACATGT TGAGCTACAG CACCAGATTC AGCAATTAAG CTCTAAGCCA	240
TCTGCAAAAA TGACCTCTTA TCAAAAGGAG CAATTAAAGG TACTCTCTAA TCCTGACCTG	300
TTGGAGTTTG CTTCGGGTCT GGTTCGCTTT GAAGCTCGAA TTAACCGCG ATATTGAAG	360
TCTTCGGGG TTCTCTTAA TCTTTTGAT GCAATCCGCT TTGCTTCTGA CTATAATAGT	420
CAGGGTAAAG ACCTGATTTT TGATTATGG TCAATCTCGT TTTCTGAAT GTTTAAAGCA	480
TTTGAGGGGG ATTCAATGAA TATTIATGAC GATTCCGAG TATTGGAGCG TATCCAGTCT	540
AAAGATTTTA CTATTACCCC CTCTGGCAAA ACTTCTTTTG CAAAAGCGCT TCGCTATTTT	600
GGTTTTTATC GTCGTCTGTT AAACGAGGGT TATGATAGTG TTGCTCTTAC TATGCTCTGT	660
AATTCCCTTT GCGGTIATGT ATCTGCATTA GTTGAATGTG GTATTCTTAA ATCTCAACTG	720
ATGAATCTTT CTACCTGTAA TAATGTTGTT CCGTIACTTC GTTTIATTAA CGTAGATTTT	780
TCCTCCCAAC GTCTGACTG CTATAATGAG CCACTTCTTA AAATCGGATA AGGTAATICA	840
CAATGATTAA AGTCAAAAT AAACCATCTC AAGCCCAATT TACTACTCGT TCTGCTGTTT	900
CTCGTCAGGG CAAGCCTTAT TCACTGAATG AGCAGCTTGG TTACGTTGAT TTGGGTAATG	960
AATATCCGGT TCTTGTCAAG ATTAAGCTTG ATGAAGGTCA GCGAGGCTAT GCGGCTAGTC	1020
TGTACACCGT TCATCTGTCC TCTTTCAAAG TTGGTCAGTT CCGTTCGCTT ATGATTGACC	1080
GTCTGCGCCT CTTTCCGGCT AAGTAACATG GAGCAGGTGG CGGATTTCGA CACAATTTAT	1140
CAGGCGATGA TAGAAATCTC CGTTGTACTT TGTTTCGGCG TTGGTATAAT CGCTGGGGGT	1200

CAAAGATGAG	TGTTTTAGTG	TATTCCTTCG	CCTCTTTCGT	TTAGGTTGG	TGCCTTCGTA	1260
GTGGCATTAC	GTATTTTACC	CGTTTAATGG	AAACTTCCTC	ATGAAAAAGT	CTTTAGTCCT	1320
CAAAGCCTCT	GTAGCCGTTG	CTACCCTCGT	TCCGATGCTG	TCTTTCCGCTG	CTGAGGGTGA	1380
CGATCCCGCA	AAAGCGGCGT	TTAACTCCCT	GCAAGCCTCA	GCGACCGAAT	ATATCGGTGA	1440
TCCGTGGCGG	ATGGTTGTTG	TCATTGTCGG	CGCAACTATC	GCTATCAAGC	TGTTTAAGAA	1500
ATTCACTCTG	AAAGCAAGCT	GATAAACCGA	TACAATTAAA	GGCTCCTTTT	GGAGCCTTNT	1560
TTTTTGGAGA	TTTTCAAGCT	GAAAAAATTA	TTATTGCGAA	TTCCTTTAGT	TGTTCTCTTC	1620
TATTCCTCACT	CGGCTGAAAC	TGTTGAAAGT	TGTTTAGCAA	AACCCCATAC	AGAAAAATCA	1680
TTTACTAAGC	TCTGGAAGA	CGACAAAAC	TTAGATCGTT	ACGGTAAC	TGAGGGTTGT	1740
CTGTGGAATG	CTACAGCGCT	TGTAGTTTCT	ACTGTCGACG	AAACTCAGTG	TTACGGTACA	1800
TGGGTTCCCTA	TTGGGCTTGC	TATCCCTGAA	AATGAGGGTG	GTGGCTCTGA	GGGTGGCGGT	1860
TCTGAGGGTG	CGGCTTCTGA	GGGTGGCGGT	ACTAAACCTC	CTGAGTACGG	TGATACACCT	1920
ATTCCGGGCT	ATACTTATAT	CAACCTCTC	GACGGCACTT	ATCCCGCTGG	TACTGACCAA	1980
AACCCCGCTA	ATCCTAATCC	TTCTCTTGAG	GAGTCTCAGC	CTCTTAATAC	TTTCATGTTT	2040
CAGAATAATA	GGTTCGAAA	TAGGCAGGGG	GCAATTAAGT	TTTATACGGG	CACCTGTTACT	2100
CAAGGCACGT	ACCCCGTTAA	AACTTATTAC	CAGTACACTC	CTGTATCATC	AAAAGCCATG	2160
TATGACGCTT	ACTGGAACGG	TAAATTCGAA	GACTGCGGCT	TCCATTCTGG	CTTTAATGAA	2220
GATCCATTCC	TTTGTAATA	TCAAGGCCAA	TGCTCTGACC	TGCCTCAACC	TGCTGTCAAT	2280
GCTGGCGGCG	GCTCTGCTGG	TGTTCTGGT	GCGGCTCTG	ACGGTGGTGG	CTCTGAGGGT	2340
GGCGGTTCTG	AGGCTGGCGG	CTCTGAGGGA	GGCGGTTCCG	GTGGTGGCTC	TGGTTCGGGT	2400
GATTTTGATT	ATGAAAAGAT	GGCAACCGCT	AATAAGGGGG	CTATGACCGA	AAATGCCGAT	2460
GAAAACGGCG	TACACTCTGA	CGCTAAAGGC	AAACTTGATT	CTGTCGCTAC	TGATTACGGT	2520
GCTGCTATCG	ATGGTTTCAT	TGGTGACGTT	TCCGGCCTTG	CTAATGGTAA	TGGTGCTACT	2580
GGTGATTTTG	CTGGCTCTAA	TTCCCAATG	GCTCAAGTGG	GTGACGGTGA	TAATTCACCT	2640
TTAATGAATA	ATTTCCGCTCA	ATATTTACCT	TCCCTCCCTC	AATCGGTTGA	ATGTCGCGCT	2700
TTTGCTTTTA	GCGCTGGTAA	ACCATATGAA	TTTTCTATTG	ATTGTGACAA	AATAAAGTTA	2760
TTCGCTGGTG	TCTTTGGGTT	TCTTTTATAT	GTGCGCACCT	TTATGTATGT	ATTTCTAGG	2820
TTTGCTAAGA	TACTCGGTAA	TAAGGAGTCT	TAATCATGCC	AGTTCTTTTG	GGTATTCGCT	2880
TATTAATGCG	TTTCCTGGT	TTCTTCTGG	TAACTTTGTT	CGGCTATCTG	CTTACTTTTC	2940
TTAAAAAGGG	CTTCGGTAAG	ATAGCTATTG	CTATTTCATT	GTTCCTTGCT	CTTATTATTG	3000
GGCTTAACTC	AATTCTGTG	GGTTATCTCT	CTGATATTAG	CGCTGAATTA	CCCTCTGACT	3060
TTGTTCAAGG	TGTTCAAGTA	ATTCTCCCGT	CTAATGGCGT	TCCCTGTTTT	TATGTTATTC	3120
TCTGTGTAAA	GGCTGCTATT	TTCAATTTTT	ACGTTAAACA	AAAAATGTTT	TCTTATTTGG	3180
ATTGGGATAA	ATAATATGGC	TGTTTATTTT	GTAAGTGGCA	AATTAGGCTC	TGGAAGACG	3240

CTCGTTAGCG	TGGTAAGAT	TCAGGATAAA	ATTGTAGCTG	GGTGGAAAA	AGCAACTAAT	3300
CTTGATTAA	GCCTTCAAAA	CCTCCCGCAA	GTGGGGAGGT	TCGCTAAAA	GCCTCGCGTT	3360
CTTAGAATAC	CGGATAAGCC	TTCTATATCT	GATTTCGCTG	CTATTGGGGC	CGGTAATGAT	3420
TCCTACGATG	AAAAAATAAA	GGGCTTGCTT	GTTCGCGATG	AGTGGGGTAC	TTGGTTTAAT	3480
ACCGGTCTCT	GGAAATGATA	GGAAAGACAG	CCGATTATTC	ATTGGTTTCT	ACATGCTCGT	3540
AAATTAGCAT	GGGATATAT	TTTTCTTGTT	CAGGACTTAT	CTATTGTGTA	TAAACAGCGG	3600
CGTTCTGCAT	TAGCTGAACA	TGTTGTTTAT	TGTCGTCGTC	TGGACAGAAT	TACTTTACCT	3660
TTTGTCGGTA	CTTTATATTC	TCTTATTACT	GGCTCGAAAA	TGCCTCTGCC	TAAATTACAT	3720
GTTCGGCTTG	TAAATATGG	CGATTCTCAA	TTAAGCCCTA	CTGTTGAGCG	TTGGCTTTAT	3780
ACTGTGAAGA	ATTGTATATA	CGCATATGAT	ACTAAACAGG	CTTTTCTAG	TAAATTATGAT	3840
TCGGGTGTTT	ATTCTTATTT	AAAGCCTTAT	TTATCACAGG	GTGGGTATTT	CAAAACCAATTA	3900
AATTTAGGTC	AGAAGATGAA	GCTTACTAAA	ATATATTGTA	AAAAGTTTTC	ACGCGTCTCT	3960
TGTCCTGCGA	TGGGATTTCG	ATCAGGATTT	ACATATAGTT	ATATAACCGA	ACCTAAGCCG	4020
GAGGTTAAAA	AGGTAGTCTC	TCAGACCTAT	GATTTTGATA	AATTCACTAT	TGACTCTTCT	4080
CAGCGCTTA	ATCTAAGCTA	TCGCTATGTT	TTCAAGGATT	CTAAGGGAAA	ATTAATTAAT	4140
AGCGACGATT	TACAGAAGCA	ACGTTATTCA	CTCAGATATA	TTGATTATG	TACTGTTTCC	4200
ATTAAAAAAG	GTAATTCAAA	TGAAAATTGTT	AAATGTAATT	AATTTTGTTT	TCTTGATGTT	4260
TGTTTCATCA	TCTTCTTTTG	CTCAGGTAAT	TGAAATGAAT	AATTCGCCCT	TGCGCGGATT	4320
TGTAAGTTGG	TATTCAAAGC	AATCAGGCGA	ATCGGTTATT	GTTCCTCCCG	ATGTAAGGAG	4380
TACTGTTACT	GTATATTTCAT	CTGACGTTAA	ACCTGAAAA	CTACGCAATT	TCTTTATTTT	4440
TGTTTACGT	GCTAATAATT	TTGATATGCT	TGCTTCAATT	CCTTCGATA	TTGAGAAATG	4500
TAATCCAAAC	AATCAGGATT	ATATTGATGA	ATTGCCATCA	TCTGATAATC	AGGAATATGA	4560
TGATAATTCC	GCTCCTTCTG	GTGGTTTCTT	TGTTCCGCAA	AATGATAATG	TTACTCAAAAC	4620
TTTTAAAAAT	AATAACGTTT	GGGCAAGGA	TTTAATACGA	GTTCGGAATT	TGTTTGTAAT	4680
GTCTAATACT	TCTAAATCCT	CAAAATGATT	ATCTATTGAC	GGTCTAATC	TATTAGTTGT	4740
TAGTGCACCT	AAAGATATTT	TAGATAACCT	TCCTCAATTC	GTTCCTACTG	TTGATTTCGC	4800
AACTGACCAG	ATATTGATTG	AGGGTTTGAT	ATTGAGGTT	CAGCAAGGTG	ATGCTTTAGA	4860
TTTTTCATTT	GCTGCTGGCT	CTCAGCGTGG	CACCTGTTGCA	GGCGGTGTTA	ATACTGACCG	4920
CCTCACCTCT	TTTTATATCT	CTGCTGGTGG	TTGTTTCCGT	ATTTTAAATG	GCGATGTTTT	4980
AGGGCTATCA	GTTCGGCAT	TAAAGACTAA	TAGCCATTCA	AAAATATTGT	CTGTGCCACG	5040
TATTCTTACG	CTTCAGGTTC	AGAAGGGTTC	TATCTCTGTT	GGCCAGAAATG	TCCCTTTTAT	5100
TACTGGTCGT	GTGACTGGTG	AATCTGCCAA	TGTAATAAAT	CCATTTCAGA	CGATTGAGCG	5160
TCAAAATGTA	GGTATTTCCA	TGAGCGTTTT	TCCTGTTGCA	ATCGCTGGCG	GTAATATTGT	5220
TCGTGATATT	ACCAGCAAGG	CCGATAGTTT	GACTTCTTCT	ACTCAGGCAA	GTGATGTTAT	5280

TACTAATCAA AGAAGTATTG CTACAACGGT TAATTGGCT GATGGACAGA CTCITTTACT	5340
CGGTGGCCTC ACTGATTATA AAAACACTTC TCAAGATTCT GCGTAGCCGT TCCTGTCTAA	5400
AATCCCTTTA ATGGGCCTCC TGTTTAGCTC CCGCTCTGAT TCCAACGAGG AAAGCACGTT	5460
ATACGTGCTC GTCAAAGCAA CCATAGTAGG CGCCCTGTAG GGGCGGATTA AGCGGGCGG	5520
CTGTGGTGGT TACGCCGAGC GTGACCGCTA CACTTGCCAG CGCCCTAGCG CCGGCTCCTT	5580
TGGCTTTCTT CCGTTCCTTT GTGCCACGTT TCGCGGCTT TCCCGGTCAA GCTCTAAATC	5640
GGGGCTCCC TTTAGGTTTC CCAITTTAGT GTTACGGCA CCTCGACCCC AAAAACTTG	5700
ATTGGGTGA TGGTTCAGCT AGTGGGCCAT CGCCCTGATA GACGGTTTTT CGCCCTTGA	5760
CGTTGGAGTC CACGTTCTTT AATAGTGGAC TCTGTTCGA AACTGGAACA ACACCTCAACC	5820
CTATCTCGG CTATTCTTTT GATTATAAG GGATTTGGC GATTTCGGAA CCACCATCAA	5880
ACAGGATTTT CGCTGTCTGG GGCAAAACAG CGTGGACCG TTGCTGCAAC TCTCTCAGGG	5940
CGAGCGGTG AAGGCAATC AGCTGTTCCC CGTCTCGTG GTGAAAAGAA AAACCAACCT	6000
GGGCCCCAAT ACGCAAAACG CCTCTCCCCG CGCGTTGGCC GATTCAATTA TGCAGCTGGC	6060
ACGACAGTTT TCCGACTGG AAAGCGGGC GTGAGCGCAA CGCAATTAAT GTGAGTTAGC	6120
TCACTCATTG GGCACCCGAG GTTTTACATT TTATGCTTCC GGCTCGTATG TTGCTGGAA	6180
TTGTAGCGG ATAACAATTT CACAGCGCTG ACTTGGCACT GCGCGTCGTT TTACAACGTC	6240
GTGACTGGGA AAACCTCGGC GTTACCCAAG GTTTGTACAT GGAGAAAAAT AAGTGAAACA	6300
AAGCACTATT CCACTGGCAC TCTTACCGTT ACTGTTTACC CCTGTGGCAA AAGCCCAGGT	6360
CCAGCTGCTC GAGTCGGTCT TCCCCTGGC ACCCTCTCC AAGAGCACTT CTGGGGGCAC	6420
AGCGGGCTG GGTGTGCTGG TCAAGACTAA TTCCCCAAC CGGTGACGGT GTCGTGGAAC	6480
TCAGGGCCCC TGACAGCGG CGTGACACCC TTCCCCGGTG TCCTAGAGTC CTCAGGACTC	6540
TACTCCCTCA GCAGCGTGGT GACCGTGCCC TCCAGCACTT TGGGCACCCA GACCTACATC	6600
TCCAACGTCA ATCAAGAGCC CAGCAACACC AAGTGGACA AGAAAGCAGA GCCCAATCT	6660
TGTACTAGTG GATCCTACCC GTACGACGTT CCGGACTACG GTTCTTAGCG TGAAGCGCAT	6720
GACCGTGCTA AGGCTGCATT CAATAGTTTA CAGGCAAGTG CTACTGAGTA GATTGGCTAC	6780
GCTTGGGCTA TGGTAGTAGT TATAGTTGGT GCTACCATAG GGATTAAATT ATTCAAAAAG	6840
TTTACGAGCA AGGCTTCTTA AGCAATAGCG AAGAGGCGCG CACCGATGCC CTTGCCAAC	6900
AGTTGGCGAG CCGTAATGGC GAATGGCGCT TTGCGTGGTT TCGGGACCA GAAGCGGTGC	6960
CGGAAAGCTG GCTGGAGTGC GATCTTCCTG AGGCGGATAC GTCGTGCTC CGCTCAAAGT	7020
GGCAGATGCA CGGTACCAT GCGGCCATCT ACACCAAGCT AACCTATCCC ATTACGGTCA	7080
ATCCGCGGTT TGTTCGCCG GAGAATCCGA CGGGTTGTTA CTCGCTCACA TTAAATGTTG	7140
ATGAAAGCTG GGTACAGGAA GGCCAGACGC GAATTATTTT TGATGGCGTT CATTATGGTT	7200
AAAAATGAG CTGATTTAAC AAAAATTTAA CGCGAATTTT AACAAAATAT TAACGTTTAC	7260
AATTAAATA TTGCTTATA CAATCTTGCT GTTTTGGGG CTTTTCTGAT TATCAACCGG	7320

GGTACATATG ATTGACATGC TAGTTTTACG ATTACCGTTC ATCGATTCTC TTGTTGCTC	7380
CAGACTCTCA GCGAATGACC TGATAGCCTT TGTAGATCTC TCAAAAATAG CTACCCCTCTC	7440
CGGCATTAAAT TTATCAGCTA GAAGCGGTGA ATATCATATT GATGGTGATT TGACTGTCTC	7500
CGGCCTTTCT CACCCCTTTG AATCTTTACC TACACATTAC TCAGGCAATG CATTAAAAAT	7560
ATATGAGGGT TCTAAAAAT TTTATCCTTG CGTTGAAATA AAGGCTTCTC CGGCAAAAGT	7620
ATTACAGGGT CATAATGTTT TTGGTACAAC CGATTTAGCT TTATGCTCTG AGGCTTTATT	7680
GCTTAATTTT GCTAATTTCT TGCCTTGCCT GTATGATTTA TTGGACGTT	7729

(2) INFORMATION FOR SEQ ID NO:4:

(1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 7557 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: both
 (D) TOPOLOGY: circular

(x1) SEQUENCE DESCRIPTION, SEQ ID NO:4:

AATGCTACTA CTATTAGTAG AATTGATGCG ACCTTTTCAG CTCGCGCGCC AAATGAAAAT	60
ATAGCTAAAC AGGTTATTGA CGATTTCGGA AATGATCTA ATGGTCAAA TAAATCTACT	120
CGTTGCGAGA ATTGGGAATC AACTGTTACA TCGAATGAAA CTTCAGACA CGTACTTTA	180
GTTGCATATT TAAACATGT TGAGCTACAG CACCAGATTC AGCAATTAA CTCTAAGCCA	240
TCGGCAAAAA TGACCTCTTA TCAAAAGGAG CAATTAAAGG TACTCTCTAA TCCTGACCTG	300
TTGGAGTTTG CTTCGGTCT GGTTCGCTTT GAAGTCCGAA TTAAAACGGC ATATTGAAG	360
TCCTTCGGGC TTCTCTTAA TCTTTTGTAT GCAATCCGCT TTGCTTCTGA CTATAATAGT	420
CAGGGTAAAG ACCTGATTTT TGATTATGG TCATTCTGCT TTTCTGAAGT GTTTAAAGCA	480
TTTGAGGGGG ATTCAATGAA TATTATGAC GATTCGGCAG TATTGGACGC TATCCAGTCT	540
AAACATTTTA CTATTACCCC CTCTGGCAAA ACTTCTTTTG CAAAAGCGTC TCGTATTTT	600
GGTTTTTATC GTGCTGTGTT AAACGAGGCT TATGATAGTG TTGCTCTTAC TATGCTCGGT	660
AATTCGTTT GCGCTTATGT ATCTGCATTA GTTGAATGTG GTATTCCTAA ATCTCAACTG	720
ATGAATCTTT CTACCTGTAA TAATGTTGTT CCGTTAGTTC GTTTTATTAA CGTAGATTTT	780
TCTTCCCAAG GTCTGACTG GTATAATGAG CGAGTTCTTA AAATGCGATA AGGTAATTCA	840
CAATGATTAA AGTTGAAATT AAACCATCTC AAGCCGAATT TACTAGTCTG TCTGTGTTT	900
CTCGTCAGGG CAAGCCTTAT TCACGTGAATG AGCAGCTTTG TTACGTGATG TTGGGTAATG	960
AATATCCGGT TCTTGTCAAG ATTACTCTTG ATGAAGGTCA GCCAGGCTAT GCGCTGGTCT	1020
TGTACACCGT TCATCTGTCC TCTTTCAAAG TTGGTCAGTT CGGTTCCTCT ATGATTGACC	1080
GTCTGGCGCT CGTTCGGGCT AAGTAACATG GAGCAGGTCG CGGATTTCGA CACAATTIAT	1140
CAGGCGATGA TACAAATCTC GGTGTGACTT TGTTCGCGC TTGGTATAAT GCGTGGGGGT	1200
CAAAGATGAG TGTTTTAGTG TATCTTTCC CCTCTTTCGT TTAGGTTGG TGCCTTCGTA	1260

GTGGCATTAC GTATTTIACC CGTTTAATGG AAACCTTCGTC ATGAAAAAGT CTTIAGTCGT	1320
CAAAGCCTCT GTACCGGTG CTACCCCTCGT TCGGATGCTG TCTTTCGGTG CTGAGGGTGA	1380
CGATCCCGCA AAAGCCGCTT TTAACCTCGT GCAAGCCTCA GCGACCGAAT ATATCGGTGA	1440
TGGTGGGGCG ATGGTTGTTG TCATTGTCGG CGCAACTATC GGTATCAAGC TGTTTAAAGAA	1500
ATTACCTCGC AAAGCAAGGT GATAAACCGA TACAATTAAA GGCTCGTTTT GAGGCTTTTT	1560
TTTTTGAGA TTTTCAACGT GAAAAATTA TTATTGCGAA TTCCTTTAGT TGTCCTTTTC	1620
TATTCTCACT CCGCTGAAAC TGTGAAAGT TGTTAGCAA AACCCGATAC AGAAAAATCA	1680
TTIACATAAC TCTGGAAAGA CQACAAAAC TTAGATCGTT ACGCTAACTA TGAGGGTTGT	1740
CTGTGAATG CTACAGCGGT TGTAGTTGT ACTGGTGACC AAACCTCAGT TTACGGTACA	1800
TGGTTCCTA TTGGGCTTGC TATCCCTGAA AATGAGGGTG GTGGCTCTGA GGGTGGGGGT	1860
TCTGAGGGTG GGGGTTCTGA GGTGCGCGT ACTAAACCTC CTGAGTACGG TGATACACCT	1920
ATTCCGGCT ATACTTATAT CAACCCCTC GACGGCACTT ATCCGCTCG TACTGAGCAA	1980
AACCCGCTA ATCTAATATC TTCTCTTAC GAGTCTCAGC CTCTTAATAC TTTCATGTTT	2040
CAGAATAATA GGTTCGAAA TAGGCAGCG GCATTAACTG TTTATACGGG CACTGTACT	2100
CAAGGCACTG ACCCGGTTAA AACTTATAC CAGTACACTC CTGTATCATC AAAAGCCATG	2160
TATGACGCTT ACTGAAACGG TAAATTGAGA TACTGCGCTT TCCATTCTCG GTTAAATGAA	2220
GATCCATTGC TTTGTGAATA TCAAGGCCAA TGTCTGACC TGCCTCAACC TCCTGTCAAT	2280
GCTGGCGGGC GCTCTGGTG TGGTCTGGT GCGCGCTCTG AGGGTGGTG CTCTGAGGGT	2340
GGCGGTTCTG AGGTGGCGG CTCTGAGGGA GGCGGTTCCG GTGGTGGCTG TGGTCCGGT	2400
GATTTTGATT ATGAAAAGAT GGCAACCGT AATAAGGGG CTATGACCGA AAATCCGGAT	2460
GAAAACGGCG TACACTCTGA CGCTAAAGGC AAACCTGATT CTGCTCGTAC TGATTACGGT	2520
GCTGCTATCG ATGGTTTCAT TGGTGACGTT TCCGCGCTTG CTAATCGTAA TGGTGCTACT	2580
GGTGATTTCG CTGGCTCTAA TTCGCAATG GCTCAAGTCG GTGACGGTGA TAAITGACCT	2640
TTAATGAATA AITTCGGTCA ATATTACCT TCCCTCGCTC AATCGGTGTA ATGCTGGCCT	2700
TTTGCTTTA GCGCTGGTAA ACCATATGAA TTTCTATTG ATGTGAGCAA AATAAACTTA	2760
TTCCGTGGTG TCTTTGGCTT TCTTTTATAT GTTCCGACCT TTAGTATGT ATTTCTACG	2820
TTTGTAACA TAGTGGCTAA TAAGGAGCTT TAATCATGCC AGTCTTTTG GGTATTCGGT	2880
TATTATTGCG TTTCCTCGGT TTGCTCTGG TAACTTTGTT CCGCTATCTG GTTACTTTTC	2940
TAAAAAGGG CTTCGGTAAG ATAGCTAATG CCTGTTTCTT GCTCTTATTA TTGGGCTTAA	3000
CTCAATCTCT GTGGGTTATC TCTCTGATAT TAGCGCTCAA TTACCGCTCG ACITTTCTGA	3060
GGGTGTTGAG TTAATTCTCC GGTCTAATGC GGTTCGGTGT TTTTATGTTA TTCTCTGTGT	3120
AAAGGCTGCT ATTTTGTATT TTGACGTTAA ACAAAAAATG GTTCTTATT TGAATTGGGA	3180
TAAATAATAT GGTGTTTTAT TTTGTAAC TGCAAAATAGG CTCTGGAAG ACGCTCGTTA	3240
GCGTTGGTAA GATTCAAGAT AAAATTGTAG CTGGGTGCAA AATAGCAACT AATCTTGATT	3300

TAAGGCTTCA	AAAGCTCCCG	CAAGTCGGGA	GGTTCGGTAA	AAGGCCTCGC	GTCTTAGAA	3360
TACCGGATAA	GCTTCTATA	TCTGATTGG	TTGCTATTGG	GCGCGGTAAT	GATTCCCTACG	3420
ATGAAAATAA	AAACGGGTTG	CTTGTTCTCG	ATGAGTGGCG	TACTTGGTTT	AATACCCGTT	3480
CTTGAATGA	TAAGGAAAGA	CAGCCGATTA	TTGATTGGTT	TCTACATGCT	CCTAAATTAG	3540
GATGGGATAT	TATTTTCTT	GTTCAGGACT	TATCTATTGT	TGATAAACAG	GCGCGTCTCG	3600
CATTAGCTGA	ACATGTTGT	TATTGTCGTC	GTCTGGACAG	AATTACTTTA	CCTTTTGTCG	3660
GTACTTTATA	TTCTCTTAT	ACTGGCTCGA	AAATGCCTCT	GCCTAAATTA	CATGTTGGCG	3720
TTGTTAAATA	TGGCGATTCT	CAATTAAAGCC	CTACTGTTGA	GCGTTGGCTT	TATACTGGTA	3780
AGAATTGTGA	TAACGCATAT	GATATAAAC	AGGCTTTTTC	TAGTAATTAT	GATTCGGGTC	3840
TTTATTCTTA	TTTAACCGCT	TATTTATCAC	ACGGTCGGTA	TTTCAAACCA	TTAAATTTAG	3900
GTGAGAAGAT	GAAGCTTACT	AAAATAATTT	TGAAAAAGTT	TTACCGGCTT	CTTTGCTTTG	3960
CGATTGGATT	TGCATCAGCA	TTACATATA	TTATATAAAC	CCAACCTAAG	CCGGAGGTTA	4020
AAAAGGTAGT	CTCTCAGACC	TATGATTTTC	ATAAATTCAC	TATTGACTCT	TCTCAGCGTC	4080
TTAATCTAAG	CTATCGCTAT	GTITTCAGGG	TATTCTAAGGG	AAAAATTAAT	AATAGCGACG	4140
ATTACAGAA	GCAAGGTTAT	TCACTCACAT	ATATTGATTT	ATGTACTGTT	TCCATTAAAA	4200
AAGCTAATTC	AAATGAAATT	GTTAAATGTA	ATTAATTTTG	TTTCTTGAT	GTTTGTTC	4260
TCATCTTCTT	TTGCTCAGGT	AATTGAAATG	AATAATTCGC	CTCTGCCGGA	TTTTGTAAT	4320
TGGTATTCAA	AGCAATCAGG	CGAATCCGTT	ATTGTTTCTC	CCGATGTAAA	AGGTACTGTT	4380
ACTGTATATT	CATCTGACGT	TAAACCTGAA	AATCTACGCA	ATTCTTTTAT	TTCTGTTTTA	4440
CGTGCTAATA	ATTTTGATAT	GTTGTTTCA	ATTCCTTCCA	TAATTCAGAA	GTATAATCCA	4500
AACAATCAGG	ATTATATTGA	TGAATTGCCA	TCATCTGATA	ATCAGGAATA	TGATGATAAT	4560
TCCGCTCCTT	CTGGTGGTTT	CTTTGTTCCG	CAAAATGATA	ATGTTACTCA	AACTTTTAAA	4620
ATTAATAACG	TTCCGCCAAA	GGATTTAATA	CGAGTTCTCG	AATTGTTTGT	AAAGCTAAT	4680
ACTTCTAAAT	CCTCAAATGT	ATTATCTAAT	GACGGCTCTA	ATCTATTAGT	TGTTAGTGCA	4740
CCTAAAGATA	TTTTAGATAA	CCTTCTCTCA	TTCTTTCTTA	CTGTTGATTT	GCCAACTGAC	4800
CAGATATTGA	TTGAGGGTTT	GATATTGAG	GTTCAGCAAG	GTGATGCTTT	AGATTTTCTA	4860
TTTGCTGCTG	GCTCTCAGCG	TGGCACTGTT	GCAGGCGGTG	TTAATACTGA	CCGCCTCAGC	4920
TCTGTTTTAT	CTTCTGCTCG	TGGTTCGTTT	GGTATTTTTA	ATGGCQATGT	TTTAGGGCTA	4980
TCAGTTCGGC	CATTAAAGAC	TAATAGCCAT	TCAAAAAATAT	TGCTGTGCGC	ACGTATTCTT	5040
ACGCTTTCAG	GTGAGAAGGG	TTCTATCTCT	GTTGGCCAGA	ATGTCCTCTT	TATTACTGGT	5100
CGTGCTAGCT	GTGAATCTGC	CAATGTAAAT	AATCCATTTC	AGACGATTGA	GCGTCAAAAT	5160
GTAGGTATTT	CCATGACCGT	TTTTCTGTTT	GCAATGGCTG	GCGGTAATAT	TGTTCTGGAT	5220
ATTACCAGCA	AGGCCGATAG	TTTCAGTTCT	TCTACTCAGG	CAAGTGATGT	TATTACTAAT	5280
CAAAGAAGTA	TTGCTACAAC	GTTAATTTG	CGTGATGGAC	AGACTCTTTT	ACTCGGTGGC	5340

CTCACTGATT ATAAAAACAC TTCTGAAGAT TCTGGCGTAC CGTTCGTGTC TAAAAATGCCCT	5400
TAAATCGGCC TCTGTTTAG CTCCCGCTCT GATTCCAACG AGGAAAGCAC GTTATACGTG	5460
CTCGTCAAG CAACCATAGT ACGGCCCCTG TAGCGGCGCA TTAAGCGCGG CGGGTGTGGT	5520
GGTTACGGCG AGCGTGACCG CTACACTTGC CAGCGCCCTA GCGCCCGCTT CTTCGCTTT	5580
CTTCGCTTCC TTCTCGCCA CGTTCGCGG CTTCGCCGT CAAGCTCTAA ATCGGGGGCT	5640
CCCTTTAGGG TTCCGATTTA GTGCTTIACG GCAGCTCGAC CCGAAAAAC TTGATTTGGG	5700
TGATGGTTCA CGTAGTGGCG CATCGCCCTG ATAGACGGTT TTTCGCCCTT TGACGTTGGA	5760
GTCCACGTTT TTTAATAGTG GACTCTTGT CCAAACTGGA ACAACACTCA ACCCTATCTC	5820
GGGCTATTCT TTGATTAT AAGGATTTT GCGGATTTCG GAACGACCAT CAAACAGGAT	5880
TTTCGCTCG TGGGCAAC CAGCGTGGAC CGCTTGCTGC AACTCTCTCA GGGCCAGGCG	5940
GTGAAGGGCA ATCAGCTGTT GCCGCTCTCG CTGGTGAAAA GAAAAACCAC CCGTGGCGCC	6000
AATACGCAA CGGCTCTCC CCGCGCGTTC GCGGATTCTA TAATGCAGCT GGCAGCAGAG	6060
GTTTCCGAC TGAAGCGG GCAGTGAGCG CAACGCAATT AATGTAGTT AGCTCACTCA	6120
TTAGGCACCC CAGGCTTAC ACTTTATGCT TCCGGCTCGT ATGTTGTGTG GAATTGTGAG	6180
CGGATAACAA TTTCACACCG CAAGGAGACA GTCTAATGA AATACCTATT GCCTACGGCA	6240
GCCGCTGGAT TGTIATTA CTGTCGCCAA CCGGCCATGG CCGAGCTCTT CCGGCCATCT	6300
GATGAGCACT TGAATCTGG AAGTCCCTCT GTTGTGTGCC TGCTGAATAA CTTCATATCC	6360
AGAGAGGCCA AAGTACAGTG GAAGTGGAT AACGCCCTCC AATCGGGTAA CTCGCCAGAG	6420
AGTGTACAG AGCAGGACAG CAAGGACAGC ACCTACAGCC TCAGCAGCAC CCTGACGGTG	6480
AGCAAAGCAG ACTACAGAA ACACAAAGTC TACGCCCTCG AAGTCAACCA TCAGGGCGCTG	6540
AGCTCGCCCG TCACAAAGAG CTTCACAGG GGAGAGTGT CTAGAACGCG TCACTTGCCA	6600
CTGGCGCTCG TTTTACAACG TCGTGACTGG GAAAACCCCTG GCGTTACCCA AGCTTAATCG	6660
CCTTGCAGAA TTCCCTTTCC CCAAGTGGCG TAATAGCGAA GAGGCCCGCA CCGATGCCCG	6720
TTCCCAACAG TTGCGGAGCC TGAATGGCGA ATGGCGCTTT GCGTGGTTTC CGGCACCAGA	6780
AGCGGTGCC GAAAGCTGGC TGGAGTGCGA TCTTCTGAG GCGATACGG TCGTCTGCCC	6840
CTCAAAGTGC CAGATGCAGG GTTACGATGC GCCCATCTAC ACCAAGCTAA CCTATCCAT	6900
TACGGTCAAT CCGCGCTTTC TTCCACCGGA GAATCCGACG GGTGTTACT CGCTCACATT	6960
TAATGTTGAT GAAAGCTGCG TACGGAAGG CCGAGCCGGA ATTATTITG ATGCGGTTCC	7020
TATTGGTTAA AAAATGAGCT GATTTAACAA AAATTTAACG CGAATTITAA CAAAATATTA	7080
ACGTTTACAA TTTAAATATT TGCTTATACA ATCTTCCGTG TTTTGGGGCT TTCTGATTA	7140
TCAACCGGGG TACATATGAT TGACATGCTA GTTTTACGAT TACCGTTCAT CGATTCTCTT	7200
GTTTGTCCA GACTCTCAGG CAATGACCTG ATAGCCTTTC TAGATCTCTC AAAAATAGCT	7260
ACGCTCTCCG GCATTAAATT ATCAGCTAGA ACGGTTGAAT ATCATATTGA TGGTGAATTG	7320
ACTGTCTCCG GCCTTCTCA CCGTTTTGAA TCTTTACCTA CACATTACTC AGGGAATTGCA	7380

TTTAAATAT ATGAGGGTTC TAAAAATTTT TATCCTTGGC TTGAAATAAA GGCTTCTCCG	7440
GCAAAAGTAT TACAGGTCAT TAATGTTTTT GGTACAACCG ATTTAGCTTT ATGCTCTGAG	7500
GCTTATTCG TTAATTTTGC TAATCTTTG CTTTGGCTGT ATGATTATT GGATGTT	7557

(2) INFORMATION FOR SEQ ID NO:5:

(1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 818 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: both
- (D) TOPOLOGY: circular

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:5:

AATGCTACTA CTATTAGTAG AATGATGCC ACCTTTTCAG CTCGCCGCCC AAATGAAAA	60
ATAGCTAAAC AGGTATTGA CGATTTCGCA ATCTATCTA ATGTCGAAAC TAAATCTACT	120
CGTTCCGACA ATTGGGAATC AACTGTTACA TCGAATGAAA CTTCCGAGACA CGGTACTTTA	180
GTTGCATATT TAAACATGT TGAGCTACAG CACGAGATTC AGCAATTAAAG CTCTAAGCCA	240
TCTGCAAAA TGACCTCTTA TCAAAAGGAG CAATTAAAGG TACTCTCTAA TCCTGACCTG	300
TTGGAGTTTG CTTCCGGTCT GGTTCGCTTT GAAGCTCGAA TTAAGGCGG ATATTGAAG	360
TCCTTCGGGC TCCTCTTAA TCTTTTGAT GCAATCGGCT TTGCTTCTGA CTATAATAGT	420
CAGGGTAAAG ACCTGATTTT TGATTATGCG TCATTCTGCT TTTCTGAAGT GTTTAAAGCA	480
TTTGAGGGGG ATTCAATGAA TATTATGAC GATTCCGCAG TATTGGACGG TATCCAGTCT	540
AAACATTTTA CTATTACCCC CTCTGGCAA ACTTCTTTTG CAAAAGGCTC TCGCTATTTT	600
GGTTTTTATC GTGCTCTGCT AAACGAGGCT TATGATAGTC TTGCTCTTAC TATGCCCTGT	660
AATTCCTTTT GCGGTTATGT ATCTGCATTA GTTGAATGTC GTATTCCTAA ATCTCAACTG	720
ATGAATCTTT CTACCTGTAA TAATGTGTT CCGTTAGTTC GTTTTATTAA CGTAGATTTT	780
TCCTCCCAAC GTCTGACTG GTATAATGAG CGAGTTCTTA AAATCGCATA AGGTAATTCA	840
CAATGATTA AGTTGAAATT AAACGATCTC AAGCCCAATT TACTAGTCGT TCTGGTGTIT	900
CTCGTCAGGG CAAGCCTTAT TCACTGAATG AGCAGCTTTG TTACGTTGAT TTGGGTAATG	960
AAATCCGGGT TCCTGTCAAG ATTACTCTTG ATGAAGGTCA GCCAGCCTAT GCGCGTGGTC	1020
TGTACACGGT TCATCTGTCC TCTTTCAAAG TTGGTCAGTT CGGTTCCCTT ATGATTGACC	1080
GTCTGCCGCT CGTTCGGGCT AAGTAACATG GAGCAGGTCG CGGATTTCGA CACAATTTAT	1140
CAGGGGATGA TACAAATCTC CGTTGTACTT TGTTCGCGC TTGCTATAAT CGCTGGGGGT	1200
CAAGATGAG TGTTTTATG TATTCTTTCG CTTCTTTCGT TTAGGTTGG TCGCTTCGTA	1260
GTGGCATTAC GTATTTTACG CGTTTAATGG AAACCTCCTC ATGAAAAAGT GTTTAGTCTT	1320
CAAGGCTCT GTAGCGGTG CTACCGCTGT TCCGATGCTG TCCTTCGCTG CTGAGGGTGA	1380
CGATCCCGCA AAAGCGGCTT TTAAGTCCCT GCAAGCCTCA GCGACCGAAT ATATCGGTTA	1440
TGGCTGGCGG ATGGTTGTTG TCATTGTCGG CGCAACTATC GGTATCAAGC TGTATAAGAA	1500

ATTCACCTCG AAAGCAAGCT GATAAACCGA TAGAATTAAA GGCTCGTTTT GGAGCCTTTT 1560
 TTTTGGAGA TTTTCAACGT GAAAAAATTA TTATTGCAA TTCCTTTAGT TGTTCTTTTC 1620
 TATTCTCACT CCGCTGAAAC TGTTGAAAGT TGTTAGCAA AACCCATAC AGAAAAATCA 1680
 TTTACTAAGC TCTGGAAAGA CACAAAAACT TTAGATCGTT ACGCTAACTA TGAGGGTTGT 1740
 CTGTGGAATG CTACAGGCGT TGTAGTTTGT ACTGCTGACG AAACCTCAGTG TTACGGTACA 1800
 TGGGTTCTTA TTGGGCTTGC TATCGCTGAA AATGAGGGTG GTGGCTCTGA GGGTGGCGGT 1860
 TCTGAGGGTG GCGGTTCTGA GGGTGGGGT ACTAAACCTC CTGAGTACGG TGATACAGCT 1920
 ATTCGGGGT ATACTTATAT CAACCCCTGC GACGGCACTT ATCGGCGTGG TACTGAGCAA 1980
 AACCCCGCTA ATCCTAATCC TTCTCTTGA GAGTCTCAGC CTCTTAATAC TTTCTAGTTT 2040
 CAGAATAATA GGTTCGAAA TAGGCAGGGG GCATTAACTG TTTATACGGG CACTGTTACT 2100
 CAAGGCACTG ACCCGGTAA AACTTATTAC CAGTACACTC CTGTATCATC AAAAGCCATG 2160
 TATGACCGTT ACTGGAACGG TAAATTCAGA GATGCGGTT TCCATTCTCG CTTTAATGAA 2220
 GATCCATTGG TTTGTGAATA TCAAGGCCAA TCGTCTGACC TGGCTCAACC TCCTGTCAAT 2280
 GCTGGCGCGC GCTCTGCTGG TGGTTCTGGT GCGGCTCTG AGGGTGGTGG CTCTGAGGGT 2340
 GCGGTTCTG AGGGTGGCGG CTCTGAGGGA GCGGTTCTG GTGGTGGCTG TGGTTCGGGT 2400
 GATTTTGATT ATGAAAAGAT GGCAACGCT AATAAGGGGG CTATGACCGA AAATGCCGAT 2460
 GAAAACGCGC TACAGTCTGA CGCTAAAGGC AAACCTIGATT CTCTCGCTAC TGATTAGGGT 2520
 GCTGCTATCG ATGGTTTCAT TGGTGACGTT TCCGGCCTTG CTAATGGTAA TGGTGCTACT 2580
 GGTGATTTTG CTGGCTCTAA TTCCCAAATG GCTCAAGTCG GTGAGGGTGA TAATTCACCT 2640
 TTAATGAATA ATTTCGCTCA ATATTIACCT TCCCTCCCTC AATCGGTTGA ATGTCGCCCT 2700
 TTTGCTTTA GCGCTGGTAA ACCATATGAA TTTTCTATTG ATTGTGACAA AATAAACTTA 2760
 TTCCGTGGTG TCTTTGCGTT TCTTTTATAT GTTCCGACCT TTATGTATCT ATTTTCTACG 2820
 TTTGCTAACA TACTGCGTAA TAAGGAGTCT TAATCATGCC AGTTCCTTTG GGTATTCCGT 2880
 TATTATTCGG TTTCTCGGT TTCTTCTGG TAACCTTGTG CCGCTATCTG CTATCTTTTC 2940
 TTA AAAAGGG GTTCGGTAAG ATAGCTATTG GTATTTCATT GTTCTTGCT CTATTATTG 3000
 GGGTAACTC AATTCTTGTG GGTATCTCT CTGATATTAG CGCTCAATTA CCGCTGACT 3060
 TTGTCAGGG TGTCAGTTA ATTCTCCGT CTAATCCGT TCCCTGTTTT TATGTTATTC 3120
 TCTCTGTAAG GCGTGTATT TTGATTTTT AGGTAAACA AAAATCGTT TCTATTTGG 3180
 ATTGGGATAA ATAATATGGC TGTTTATTTT GTAACGGCA AATTAGGGTC TGGAAAGAGC 3240
 CTCGTAGCG TTGTAAGAT TCAGGATAAA ATTGTAGCTG GGTGCAAAAT AGCAACTAAT 3300
 CTTGATTTAA GCGTTCAAAA CCTCCCGCAA GTCCGGAGGT TCGCTAAAAA GCGTCCGGT 3360
 CTTAGAATAC CGATAAGCC TTCTATATCT GATTTCGTTG CTATTGGGGC CGGTAATGAT 3420
 TCCTACGATG AAAATAAAAA CGGCTTGCTT GTTCTCGATG AGTCCGGTAC TTGCTTTAAT 3480
 ACCCGTCTCT GGAATGATAA GGAAGACAG CCGATTATTG ATTGTTTCT ACATGCTCGT 3540

AAATTAGGAT GGGATATTAT TTTCTTGTT CAGGACTTAT CTATTGTTGA TAAACAGCGG	3600
CGTCTGCAT TAGCTGAACA TGTGTTTAT TGTGCTGTC TGCACAGAAT TACTTTACCT	3660
TTTGTGGTA CTTTATATTC TCTTATTACT GCGTCGAAA TCGCTCTGCC TAAATTACAT	3720
GTTGGCGTGG TAAATATGG CGATTCTCAA TTAAGCCCTA CTGTTGAGCG TTGCGTTTAT	3780
ACTGGTAAGA ATTTGTATAA GGCATATGAT ACTAAACAGG CTTTTCTAG TAATTATGAT	3840
TCCGCTGTTT ATCTTATTT AACCGCTTAT TTATCACAAG GTCCGTATTT CAAACCATTA	3900
AAATTAGGTC AGAAGATGAA GCTTACTAAA ATATATTTGA AAAAGTTTTC AGCGGTCTCT	3960
TGCTTTCGA TTGGATTTC ATCAGCATTT ACATATAGTT ATATAACCGA ACCTAAGCGG	4020
GAGGTTAAAA AGGTAGTCTC TCAGACCTAT GATTTTGATA AATTCACATAT TGACTCTTCT	4080
CAGCGTGTTA ATCTAAGCTA TCGCTATGTT TTCAAGGATT CTAAGGGAAA ATTAAATTAAT	4140
AGCGCAGATT TACAGAAGCA AGGTTATTC CTCACATATA TTGATTTATG TACTGTTTCC	4200
ATTAAAAAG GTAATCAAA TGAAATGCTT AAATGTAATT AATTTTGTTT TCTTGATGTT	4260
TGTTTCATCA TCTTCTTTG CTCAGGTAAT TGAAATGAAT AATTCGCGTC TCGCGCATTT	4320
TGTAACCTGG TATTCAAAGC AATCAGCGCA ATCCGTTATT GTTCTCCCG ATGTAAAAGG	4380
TACTGTTACT GTATATTCAT CTCAGCTTAA ACTGAAAAAT CTACGCAATT TCTTTATTTT	4440
TGTTTTACGT GCTAATAATT TTGATATGTT TGGTTCAATT CCTCCATAA TTCAGAAGTA	4500
TAATCCGAAC AATCAGGATT ATATTGATGA ATTCCCATCA CTCGATAATC AGGAATATGA	4560
TGATAATTCC GCTCCTCTG GTGGTTTCTT TGTTCGGCAA AATGATAATG TTACTCAAAC	4620
TTTTAAAAAT AATAACGTTT GGGCAAAAGG TTTAATACCA GTTCTCGAAT TGTTTGTAAA	4680
GTCTAATACT TCTAAATCCT CAAATGTATT ATCTATTGAC GGCTCTAATC TATTAGTTGT	4740
TAGTCCACCT AAAGATATTT TAGATAACCT TCCTCAATTC CTTTCTACTG TTGATTTGCC	4800
AACTGACCAG ATATTGATTG AGGGTTTGAT ATTTGAGGTT CAGCAAGGTG ATGCTTTAGA	4860
TTTTTCATTT GCTGCTGGCT CTCAGCGTGG CACTGTTCGA GCGGCTGTTA ATACTGACCG	4920
CCTCACCTCT GTTTTATCTT CTGCTGGTGG TTCGTTGGGT ATTTTAAATG GCGATGTTTT	4980
AGGGCTATCA GTTCGGGCAT TAAAGACTAA TAGCCATTCA AAATATTGTT CTGTGCCAGG	5040
TATTCATTAG GTTCAGGTC AGAAGGTTTC TATCTCTGTT GGCAGAAATG TCCCTTTTAT	5100
TACTGCTGCT GTGACTGGTG AATCTGCCAA TGTAAATAAT CGATTTCAGA CGATTGACGG	5160
TCAAAATGTA GGTATTTCCA TGAGCGTTTT TCTGTTCCA ATCGGTGGCG GTAATATTGT	5220
TCTGATATTT ACCAGCAAGG CCGATAGTTT GAGTTCTTCT ACTCAGGCAA GTGATGTTAT	5280
TACTAATCAA AGAAGTATTG CTACAACGGT TAAITTCGGT GATGGACAGA CTGTTTTACT	5340
CGGTGGCCTC ACTGATTATA AAAACACTTC TCAAGATTCT GCGCTACGGT TCCTGCTTAA	5400
AATCGCTTTA ATCGGCTCC TGTTTAGGTC CCGCTGTGAT TCCAAGCAGG AAAGCACGTT	5460
ATACGTGCTC GTCAAGGCAA CCATAGTAGG CGCCCTGTAG CGGCGCATTA AGCGCGGCGG	5520
GTGTGGTGGT TACGCGCAGG GTGACCGCTA CACTTGCCAG GCGCCTAGCG CCGCTCTCTT	5580

TCGCTTTCCT	CGCTTCCTTT	CTCGCCACGT	TCGCGGGCTT	TCGCGGTCAA	GCTCTAAATC	5640
GGGGGCTCCC	TTTAGGGTTC	CGATTTAGTG	CTTTACGGCA	CCTCGACCCC	AAAAAACTTG	5700
ATTTGGGTGA	TGGTTCACGT	AGTGGGCAT	CGCCTGATA	GACGGTTTTT	CGCCCTTTGA	5760
CGTTGGAGTC	CACGTTCITT	AATACTGGAC	TCTTGTTCGA	AACTGGAACA	ACACTCAACC	5820
CTATCTCGGG	CTATCTTTTT	GATTTATAAG	GGATTTTGCC	GATTTGCGAA	CCACCATCAA	5880
ACAGGATTTT	CGCCTGCTGG	GGCAAAACAG	CGTGACCCGC	TTGCTGCAAC	TCCTCAGGG	5940
CCAGGCGGTG	AAGGCAATC	AGCTGTTECC	CGTCTCGCTG	GTGAAAAGAA	AAACCACCCT	6000
GGCGCCCAAT	ACGCAAAACG	CCTCTCCCGG	CGCGTTGGCC	GATTCATTAA	TGCAGCTGGC	6060
ACGACAGGTT	TCCCGACTGG	AAAGCGGGCA	GTGAGCGCAA	CGCAATTAAT	GTGAGTTAGC	6120
TCACTCATTAA	GGCACCACAG	GCTTTACACT	TTATGCTTCC	GGCTCGTATG	TTGTGTGGAA	6180
TTGTGAGCGG	ATAACAATTT	CACACGCCAA	CGAGACAGTC	ATAATGAAAT	ACCTATTGCC	6240
TACGGCAGCC	CGTGATTCT	TATTACTCGG	TGCCCCAACCA	GCCATGGCCG	AGCTGTTCCC	6300
GCCATCTGAT	GAGCAGTGA	AATCTGGAAC	TGCTCTGTGT	GTGTGCTCTG	TGAATAACTT	6360
CTATCCAGCA	GAGGCCAAAG	TACACTGGAA	GGTGCATAAC	GCCCTCCAAT	CGGCTAACTC	6420
CGAGGAGAGT	GTACACAGAG	AGGACAGCAA	GGACAGCACC	TACAGCCTCA	GCAGCACCTT	6480
GACGCTGAGC	AAAGCAGACT	ACGAGAAACA	CAAAAGTCTAC	GCCTGCGAAG	TACCCGATCA	6540
GGGCTGAGC	TGCGCCGTCA	CAAAAGAGCTT	CAACAGGGGA	GAGTGTCTTA	GAACCCGTCA	6600
CTTGGCACTG	GCGCTCGTTT	TACAACGCTG	TGACTGGGAA	AACCTTGGCG	TTACCCAAAGC	6660
TTGTACATG	GAGAAAAATA	AGTGAAACAA	AGCACTATTG	CACTGGCACT	CTTACCGTTA	6720
CTGTTTACCC	CTGTGGCAAA	AGCCGCTCC	ACCAAGGGCG	CATCGGTCTT	CCCCCTGGCA	6780
CCCTCTCCA	AGAGCACCTC	TGGGGGCACA	GCGGCCCTGG	GCTGCTGGT	CAAGACTAAT	6840
TCCCGGAACC	GGTGACGGTG	TGCTGGAAT	CAGGCGCCCT	CACCAGGGGC	GTGCACACCT	6900
TCCCGGTGT	CCTACAGTCC	TCAGGACTCT	ACTCCCTCAG	CAGCGTGGTG	ACCGTGCCCT	6960
CGAGCAGCTT	GGGACCCAG	ACCTACATCT	GCAACGTGAA	TCAAAAGCCC	AGCAACACCA	7020
AGGTGACAA	GAAAGCAGAG	CCCAAACTTT	GTACTAGTGG	ATCGTACCCG	TACGACGTTT	7080
CGGACTAGGC	TTCTTAGGCT	GAAAGCGATG	ACCCTGCTAA	GGCTGCAATG	AATAGTTTAC	7140
AGGCAAGTGC	TACTGAGTAC	ATTGCTTACG	CTTGGGCTAT	GGTAGTAGTT	ATAGTTGGTG	7200
CTACCATAGG	GATTAATAAT	TTCAAAAAGT	TTACGAGCAA	GGCTTCTTAA	GCAATAGCGA	7260
AGAGGCCCGC	ACCGATCGCC	CTTCCCAACA	GTTGCGGAGC	CTGAATGGCG	AATGGCGGTT	7320
TGCTGTGTTT	CGGCGACCAG	AAGCGGTGCC	GGAAGCTGG	CTGGAGTGG	ATCTTCTGTA	7380
GGCGGATACG	GTGCTGTGCC	CCTCAAACTG	GCAGATGCAC	GGTTACGATG	CGCCGATCTA	7440
CACCAACGTA	ACCTATCCCA	TTACGGTCAA	TCCGCGGTTT	GTTCCGACGG	GAATCCGAC	7500
GGTTGTGTAC	TGCTCACAAT	TTAATGTTGA	TGAAAGCTGG	CTACAGGAAG	GCCAGACGCG	7560
AATTAATTTT	GATGGCGTTC	CTATTGGTTA	AAAAATGAGC	TGATTTAACA	AAATTTTAAC	7620

GGCAATTTTA ACAAAATATT AACGTTTACA ATTTAAATAT TTGCTTATAC AATCTTCTG	7680
TTTTTGGGGG TTTTCGATT ATCAACCGGG GTACATATGA TTGACATGCT AGTTTACGA	7740
TTACCGTICA TCGATTCTCT TGTTCCTCC AGACTCTCAG GCAATGACCT GATAGCCTTT	7800
GTAGATCTCT CAAAAATAGC TACCCTCTCC GGCATTAATT TATCAGCTAG AACGGTTGAA	7860
TATCATATTG ATGGTGATT GACTGTCTCC GGCCTTCTC ACCCTTTTGA ATCTTACCT	7920
ACACATTACT CAGGCATTTC ATTTAAATA TATGAGGGT CTAATAATT TATCCTTGC	7980
GTTGAAATAA AGGCTTCTCC CGCAAAAGTA TTACAGGGTC ATAATGTTT TGGTACAACC	8040
GATTAGCTT TATGCTCTGA GCGTTTATTG CTAAATTTG CTAATCTTT GCCTTGCCTC	8100
TATGATTIAT TGGACGTT	8118

(2) INFORMATION FOR SEQ ID NO:6:

(1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 22 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ix) FEATURE:

- (A) NAME/KEY: misc_difference
- (B) LOCATION: replace(5, "")
- (D) OTHER INFORMATION: /note= "S REPRESENTS EQUAL MIXTURE OF G AND C"

(ix) FEATURE:

- (A) NAME/KEY: misc_difference
- (B) LOCATION: replace(6, "")
- (D) OTHER INFORMATION: /note= "M REPRESENTS EQUAL MIXTURE OF A AND C"

(ix) FEATURE:

- (A) NAME/KEY: misc_difference
- (B) LOCATION: replace(8, "")
- (D) OTHER INFORMATION: /note= "R REPRESENTS EQUAL MIXTURE OF A AND C"

(ix) FEATURE:

- (A) NAME/KEY: misc_difference
- (B) LOCATION: replace(11, "")
- (D) OTHER INFORMATION: /note= "K REPRESENTS EQUAL MIXTURE OF G AND T"

(ix) FEATURE:

- (A) NAME/KEY: misc_difference
- (B) LOCATION: replace(20, "")
- (D) OTHER INFORMATION: /note= "W REPRESENTS EQUAL MIXTURE OF A AND T"

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:6:

AGGTSMARCT KCTCGAGTCW GG

(2) INFORMATION FOR SEQ ID NO:7:

- (1) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

AGGTCAGCT GCTCGAGTCT GG

22

(2) INFORMATION FOR SEQ ID NO:8:

- (1) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

AGGTCAGCT GCTCGAGTCA GG

22

(2) INFORMATION FOR SEQ ID NO:9:

- (1) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

AGGTCAGCT TCTCGAGTCT GG

22

(2) INFORMATION FOR SEQ ID NO:10:

- (1) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

AGGTCAGCT TCTCGAGTCA GG

22

(2) INFORMATION FOR SEQ ID NO:11:

- (1) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:11:

AGGTCCAAC TCTCGAGTCT GG

22

(2) INFORMATION FOR SEQ ID NO:12:

- (1) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:12:

AGGTCCAAC TCTCGAGTCA GG

22

(2) INFORMATION FOR SEQ ID NO:13:

- (1) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:13:

AGGTCCAAC TCTCGAGTCT GG

22

(2) INFORMATION FOR SEQ ID NO:14:

- (1) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:14:

AGGTCCAAC TCTCGAGTCA GG

22

(2) INFORMATION FOR SEQ ID NO:15:

- (1) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ix) FEATURE:

- (A) NAME/KEY: misc_difference
(B) LOCATION: replace(5..6, "")
(D) OTHER INFORMATION: /note= "N-INOSINE"

(ix) FEATURE:

- (A) NAME/KEY: misc_difference
(B) LOCATION: replace(8, "")
(D) OTHER INFORMATION: /note= "N-INOSINE"

(ix) FEATURE:
 (A) NAME/KEY: misc difference
 (B) LOCATION: replace(11, "")
 (D) OTHER INFORMATION: /note= "N-INOSINE"

(ix) FEATURE:
 (A) NAME/KEY: misc difference
 (B) LOCATION: replace(20, "")
 (D) OTHER INFORMATION: /note= "W REPRESENTS EQUAL MIXTURE OF A AND T"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

AGGTNNANCT NCTCGAGTCW GG

22

(2) INFORMATION FOR SEQ ID NO:16:

(1) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 38 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

CTATTAACCTA GTAACGGTAA CAGTGGTGCC TTGCCCGA

38

(2) INFORMATION FOR SEQ ID NO:17:

(1) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 30 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

AGGCCTACTA GTACAATCCC TGGGCACAAT

30

(2) INFORMATION FOR SEQ ID NO:18:

(1) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 32 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

CCAGTTCGGA GCTCGTTGTG ACTCAGGAAT CT

32

(2) INFORMATION FOR SEQ ID NO:19:

(1) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 32 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

- (x1) SEQUENCE DESCRIPTION: SEQ ID NO:19:
CCAGTTCGGA GCTCTGTG ACAGAGCGCG CC 32
- (2) INFORMATION FOR SEQ ID NO:20:
(1) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 32 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
- (x1) SEQUENCE DESCRIPTION: SEQ ID NO:20:
CCAGTTCGGA GTCGTGCTC ACCAGTCTC CA 32
- (2) INFORMATION FOR SEQ ID NO:21:
(1) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 32 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
- (x1) SEQUENCE DESCRIPTION: SEQ ID NO:21:
CCAGTTCGGA GTCAGATG ACCAGTCTC CA 32
- (2) INFORMATION FOR SEQ ID NO:22:
(1) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 32 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
- (x1) SEQUENCE DESCRIPTION: SEQ ID NO:22:
CCAGATGTA GTCGTGATG ACCAGACTC CA 32
- (2) INFORMATION FOR SEQ ID NO:23:
(1) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 32 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
- (x1) SEQUENCE DESCRIPTION: SEQ ID NO:23:
CCAGATGTA GTCGTGATG ACCAGTCTC CA 32
- (2) INFORMATION FOR SEQ ID NO:24:
(1) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 32 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

- (x1) SEQUENCE DESCRIPTION: SEQ ID NO:24:
CCAGTTCGA GCTCGTGATC ACACAGTCTC CA 32
- (2) INFORMATION FOR SEQ ID NO:25:
- (1) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 32 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
- (x1) SEQUENCE DESCRIPTION: SEQ ID NO:25:
GCAGATTCT AGAGTTTCAG CTCAGCTTG CC 32
- (2) INFORMATION FOR SEQ ID NO:26:
- (1) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 34 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
- (x1) SEQUENCE DESCRIPTION: SEQ ID NO:26:
GCGCGTCTA GAATTACAC TCATTCTGT TGAA 34
- (2) INFORMATION FOR SEQ ID NO:27:
- (1) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 37 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
- (x1) SEQUENCE DESCRIPTION: SEQ ID NO:27:
GATCCTAGGC TGAAGCGCAT CACCTGCTA AGGCTGC 37
- (2) INFORMATION FOR SEQ ID NO:28:
- (1) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 35 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
- (x1) SEQUENCE DESCRIPTION: SEQ ID NO:28:
ATTCAATAGT TTACAGGCAA GTGCTACTGA GTACA 35

(2) INFORMATION FOR SEQ ID NO:29:

- (1) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 35 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:29:

ITGGGTACGG TTGGGCTATG GTAGTAGTTA TAGTT

35

(2) INFORMATION FOR SEQ ID NO:30:

- (1) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 35 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:30:

GGTGCTACCA TAGGGATTAA ATTATTCAAA AAGTT

35

(2) INFORMATION FOR SEQ ID NO:31:

- (1) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 18 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:31:

TACGAGCAAG GCTTCTTA

18

(2) INFORMATION FOR SEQ ID NO:32:

- (1) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 39 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:32:

AGCTTAAGAA GCCTTGCTCG TAAAGTTTTT GAATAATTT

39

(2) INFORMATION FOR SEQ ID NO:33:

- (1) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 36 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

- (x1) SEQUENCE DESCRIPTION: SEQ ID NO:33:
AATCCCTATG GTAGCAOCAA CTATAACTAC TACCAT 36
- (2) INFORMATION FOR SEQ ID NO:34:
- (1) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 35 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
- (x1) SEQUENCE DESCRIPTION: SEQ ID NO:34:
AGCCCAAGCG TAGGCAATGT ACTCACTAGC ACTTG 35
- (2) INFORMATION FOR SEQ ID NO:35:
- (1) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 34 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
- (x1) SEQUENCE DESCRIPTION: SEQ ID NO:35:
CCTGTAAACT ATTGAATGCA GCCTTAGCAG GGTC 34
- (2) INFORMATION FOR SEQ ID NO:36:
- (1) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 16 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
- (x1) SEQUENCE DESCRIPTION: SEQ ID NO:36:
ATGCCCTTCA GCCTAG 16
- (2) INFORMATION FOR SEQ ID NO:37:
- (1) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
- (x1) SEQUENCE DESCRIPTION: SEQ ID NO:37:
CATTTTGGCA GATGGCTTAG A 21
- (2) INFORMATION FOR SEQ ID NO:38:
- (1) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 18 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:38:
TAGCATTAAAC GTCCAATA 18

(2) INFORMATION FOR SEQ ID NO:39:

(1) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 26 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:39:
ATATATTITTA GTAAGCTTCA TCTTCT 26

(2) INFORMATION FOR SEQ ID NO:40:

(1) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 23 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:40:
GACAAAGAAC CGGTGAAAAAC TTT 23

(2) INFORMATION FOR SEQ ID NO:41:

(1) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 35 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:41:
GCGGGCCTCT TCGGTATTGC TTAAGAAGCC TTGCT 35

(2) INFORMATION FOR SEQ ID NO:42:

(1) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 43 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:42:
AAACGACGGC CAGTGCCAAG TGACGCGTGT GAAATTGTTA TCC 43

(2) INFORMATION FOR SEQ ID NO:43:

- (1) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 43 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:43:

GGCGAAAGGG AATTCTGCAA GCGGATTAAG CTGGGTAAC GCC

43

(2) INFORMATION FOR SEQ ID NO:44:

- (1) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 36 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:44:

GGCGTTACCC AAGCTTTGTA CATGGAGAAA ATAAAG

36

(2) INFORMATION FOR SEQ ID NO:45:

- (1) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 42 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:45:

TGAACAAAAG CACTATTGCA CTGGCACTCT TACCGTTACC GT

42

(2) INFORMATION FOR SEQ ID NO:46:

- (1) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 42 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:46:

TACTGTTIAC CCCTGTGACA AAAGCGGCCC AGTCCAGCT GC

42

(2) INFORMATION FOR SEQ ID NO:47:

- (1) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 44 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:47:
TCGAGTCAGG CCTATTGTC CCAGGGATTG TACTAGTGGA TCGG

44

(2) INFORMATION FOR SEQ ID NO:48:

- (1) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 38 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:48:
TGGCGAAAGG GAATTCGGAT CCACTAGTAC AATCCGTC

38

(2) INFORMATION FOR SEQ ID NO:49:

- (1) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 42 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:49:
GGCACAATAG GCCTGACTCG AGCAGCTGGA CCAGGCGGCG TT

42

(2) INFORMATION FOR SEQ ID NO:50:

- (1) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 42 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:50:
TGTTCACAGG GCTAAACAGT AACGGTAACG GTAAGTGTCG CA

42

(2) INFORMATION FOR SEQ ID NO:51:

- (1) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 42 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:51:
GTGCAATAGT GCTTGTTC ACTTATTTT CTCCATGTAC AA

42

(2) INFORMATION FOR SEQ ID NO:52:

- (1) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

- (x1) SEQUENCE DESCRIPTION: SEQ ID NO:52:
TAACGGTAAG AGTGGCACTG C 21
- (2) INFORMATION FOR SEQ ID NO:53:
(1) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 32 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
- (x1) SEQUENCE DESCRIPTION: SEQ ID NO:53:
CACCTTCATG AATTCGGCAA GGAGACAGTC AT 32
- (2) INFORMATION FOR SEQ ID NO:54:
(1) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
- (x1) SEQUENCE DESCRIPTION: SEQ ID NO:54:
AATTCGGCAA GGAGACAGTC AT 22
- (2) INFORMATION FOR SEQ ID NO:55:
(1) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 39 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
- (x1) SEQUENCE DESCRIPTION: SEQ ID NO:56:
AATGAAATAC GTATTGCCA CGGAGCCGC TCGATTGT 39
- (2) INFORMATION FOR SEQ ID NO:56:
(1) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 39 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
- (x1) SEQUENCE DESCRIPTION: SEQ ID NO:56:
ATTACTCGGT GCCCAACCAG CCATGGCCGA GCTCCTGAT 39

(2) INFORMATION FOR SEQ ID NO:57:

- (1) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 39 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:57:

GACCCAGACT CCAGATATCC AAGACGAATG AGTGTTAAT

39

(2) INFORMATION FOR SEQ ID NO:58:

- (1) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 13 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:58:

TCTAGAACGC GTC

13

(2) INFORMATION FOR SEQ ID NO:59:

- (1) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 45 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:59:

TTCAGGTTGA AGCTTACCGG TTCTAGAATT AACACTCAAT CCTGT

45

(2) INFORMATION FOR SEQ ID NO:60:

- (1) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 39 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:60:

TGGATATCTG GAGTCTGGGT CATCAGGAGC TCGGCCATG

39

(2) INFORMATION FOR SEQ ID NO:61:

- (1) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 39 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

- (x1) SEQUENCE DESCRIPTION: SEQ ID NO:61:
GCTGGTTGGG CAGCGATTAA TAACAATCCA GCGGCTGCC 39
- (2) INFORMATION FOR SEQ ID NO:62:
- (1) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 37 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
- (x1) SEQUENCE DESCRIPTION: SEQ ID NO:62:
GTAGGCAATA GGTATTTCAT TATGCTGTC CTTGGCG 37
- (2) INFORMATION FOR SEQ ID NO:63:
- (1) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
- (x1) SEQUENCE DESCRIPTION: SEQ ID NO:63:
TGACTGTCTC CTTGGCGTGT GAAATTGTTA 30
- (2) INFORMATION FOR SEQ ID NO:64:
- (1) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 36 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
- (x1) SEQUENCE DESCRIPTION: SEQ ID NO:64:
TAACACTCAT TCCGGATGGA ATTCTGGAGT CTGGGT 36
- (2) INFORMATION FOR SEQ ID NO:65:
- (1) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
- (x1) SEQUENCE DESCRIPTION: SEQ ID NO:65:
GCCAGTGCCA AGTCAGCGGT TCTA 24
- (2) INFORMATION FOR SEQ ID NO:66:
- (1) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 26 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:66:

ATATATTTTA GTAAGCTTCA TCTTCT

26

(2) INFORMATION FOR SEQ ID NO:67:

- (1) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 23 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:67:

GACAAAGAAC GCGTCAAAAC TTT

23

(2) INFORMATION FOR SEQ ID NO:68:

- (1) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 76 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:68:

CTGAACCTGT CTGGGACCAC AGTTGATGCT ATAGGATCAG ATCTAGAATT CATTAGAGA

60

CTGGCCTGGC TTCTGC

76

(2) INFORMATION FOR SEQ ID NO:69:

- (1) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 80 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:69:

TCGACCGTIG GTAGGAATAA TCGAATTAAT GGAGTAGCTC TAAATTCAGA ATTCATCTAC

60

ACCCAGTCCA TCCAGTAGCT

80

(2) INFORMATION FOR SEQ ID NO:70:

- (1) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 27 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:70:

GGTAAACAGT AACGGTAAGA GTGCCAG

27

(2) INFORMATION FOR SEQ ID NO:71:

- (1) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 54 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:71:

CGCCTTCAGC CTAAGAAGCG TAGTCCGGAA CGTCGTACGG GTAGGATCCA CTAG

54

(2) INFORMATION FOR SEQ ID NO:72:

- (1) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 41 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:72:

CAGCGGTTCG GGAATTAGT CITGACCAGG CAGCCAGGG C

41

(2) INFORMATION FOR SEQ ID NO:73:

- (1) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 51 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:73:

ATTCCACACA TTATACGAGC CGGAAGCATA AAGTGTCAAG CCTGGGGTGC C

51

(2) INFORMATION FOR SEQ ID NO:74:

- (1) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 42 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:74:

CTGCTCATCA GATGGCGGA AGAGTCGGC CATGGCTGCT TC

42

(2) INFORMATION FOR SEQ ID NO:75:

- (1) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 42 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

78

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:75:
GAACAGAGTC ACCGAGGGGG ~~CGAGCTGGG~~ CATGGCTGGT TG